Vol. 48 FEBRUARY, 1957 No. 2

The Journal of Educational Psychology

UNIVERSITY OF MICHIGAN
JUN
PERIODICAL READING ROOM

CONTENTS

A Second Quarter Century of Delayed Recall or Relearning at 80... 65
L. B. HILL

Recruiting Prospective Rural Teachers at the Men's Teachers' College, Tripoli, Libya..... 69
IBRAHIM E. METAWEH

The Intelligence Test in School Use: Some Persistent Issues..... 77
J. B. STROUD

The Influence of Intention on the Retention of School Materials.... 87
DAVID P. AUSUBEL, SEYMOUR H. SCHPOONT, AND LILLIAN CUKIER

The Factorial Invariance of Pure-Factor Tests..... 93
JOHN W. FRENCH

Retention of General Psychology..... 110
W. J. McKEACHIE AND DANIEL SOLOMON

Assessment of Intelligence Level by the Rorschach Test as Compared with Objective Tests..... 113
BOLESLAW A. WYSOCKI

Book Reviews..... 118

Publications Received..... 126

Published monthly except June to September
$7.00 a year in U. S. and Pan America; Canada, $7.20; other countries, $7.40.
Single issues $1.20

WARWICK & YORK, INC.
BALTIMORE 2, MD.

Entered as Second Class Matter Nov. 15, 1921, at the Post Office at Baltimore, Md. under the Act of March 3, 1879.

THE JOURNAL OF Educational Psychology

Established 1910

BOARD OF EDITORS*

WM. CLARK TROW, *Chairman*
University of Michigan

STEPHEN M. COREY
Teachers College
Columbia University

J. B. STROUD
State University of Iowa

*Appointed by the American Psychological Association.

EDITORIAL CONSULTANTS

PAUL BLOMMERS
CHESTER W. HARRIS
HAROLD E. JONES
H. H. REMMERS
PERCIVAL M. SYMONDS
MILES A. TINKER
ALEXANDER G. WESMAN
PAUL A. WITTY

THE JOURNAL OF EDUCATIONAL PSYCHOLOGY is devoted primarily to the scientific study of problems of learning, teaching, and measurement of the psychological development of the individual. THE JOURNAL will contain articles on the following subjects: the psychology of school subjects; experimental studies of learning; the development of interests, attitudes, and personality, particularly as related to school adjustment; emotion, motivation, and character; mental development and methods. This last will include tests, statistical techniques, and research techniques in cross-sectional and developmental studies.

INFORMATION FOR CONTRIBUTORS

Manuscripts and communications regarding editorial matters should be addressed to the Chairman of the Board of Editors, Wm. Clark Trow, School of Education, University of Michigan, Ann Arbor, Mich.

THE JOURNAL has set regulations regarding content and style of material published, and these should be observed in the preparation of manuscripts to be submitted.

Tables and graphs.—In keeping with the practice of most professional periodicals of like character, THE JOURNAL now requires authors to bear part of the added cost resulting from the use of tables, formulas, and graphs. Prospective contributors may obtain detailed statement as to

(Continued on Inside Back Cover)

THE JOURNAL OF
EDUCATIONAL PSYCHOLOGY

Volume 48 February, 1957 Number 2

A SECOND QUARTER CENTURY OF DELAYED RECALL OR RELEARNING AT 80

L. B. HILL

Fifty years ago, A. E. Rejall and L. B. Hill (the present writer, referred to later in this study as the experimenter), then students at Teachers' College, Columbia University, learned to typewrite with a certain degree of proficiency, under experimental conditions. The experiment consisted of learning to typewrite by writing a one hundred-word paragraph, and by copying, at the same sitting, a page of new material containing three hundred words. The original data were analyzed by the late Dr. Edward L. Thorndike and published in the *Pedagogical Seminary* of December, 1913, under the title "Practice in the Case of Typewriting."[1]

Twenty-five years later, the writer, who had had very little significant practice in using the typewriter in the meantime, recalled the learning formerly acquired. The method of recall was the same as that used in the earlier experiment. The data thus secured were published in the *Pedagogical Seminary and Journal of Genetic Psychology* under the title "A Quarter Century of Delayed Recall."[2]

Now, twenty-five years after the first recall, and fifty years after the first learning, the writer presents herewith certain data from a second quarter century of delayed recall, together with his own interpretation and comments upon the findings.

Certain facts, more or less significant to a second recall, should be kept in mind, such as the age of the learner and the fact that he retired from teaching ten years ago. Since retirement he has done no significant work of a professional nature. He has changed en-

[1] Hill, L. B., A. E. Rejall, and E. L. Thorndike. Practice in the Case of Typewriting. *Ped. Sem.*, 1913, 20, 315–529.

Hill, L. B. A Quarter Century of Delayed Recall. *Ped. Sem. and Jo. Gen. Psy.* 1934, 44, 231–238.

vironment and has been identified with some civic organizations in the community in which he has lived. He has engaged in some pleasurable activities such as taking care of a new home, planting flowers and shrubbery, and learning to play shuffleboard, to dance, and to play a better game of bridge. His physical condition may be described as good. The only real physical impairment which he recognizes is in eyesight which makes his reading a bit slow and difficult.

The second recall of learning to typewrite was undertaken (1) because of a persistent interest in the learning process—an interest which lends pleasure in learning anything new, even at the expense of considerable effort. (2) Since writing with a pen has become a little difficult, it will be worthwhile and pleasurable to be able to use the typewriter. (3) It will also be interesting to find out whether the experimenter, at his present age, can still learn, and to see what factors enter into the process. Other factors favorable to undertaking the project may be mentioned as a feeling of well-being, living in an environment where a great many old people learn new hobbies, play a great deal, and enjoy life in retirement.

Reference to Figure 1 shows that in the first experiment the learner wrote the one hundred-word paragraph at the rate of 7.4 words per minute at the first sitting. (In all three experiments two

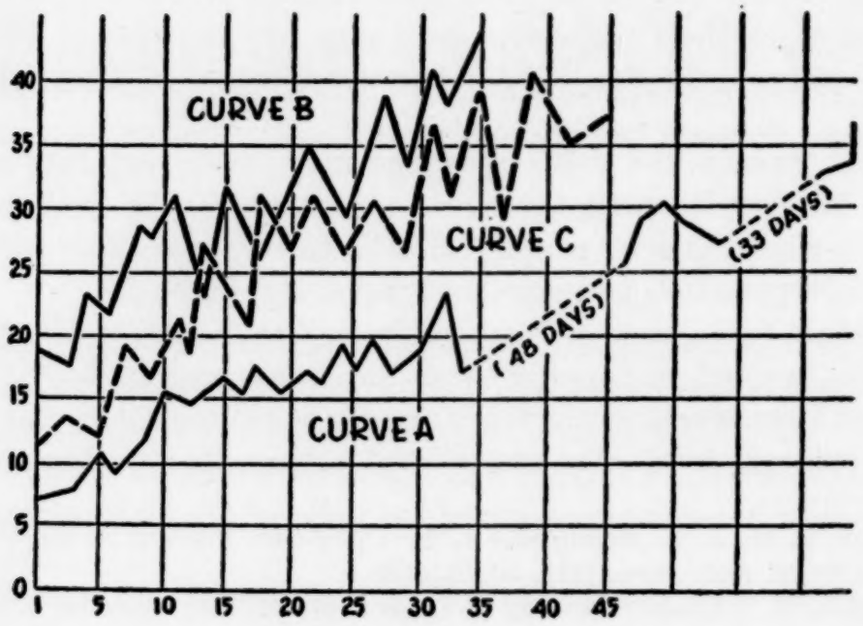

FIG. 1

words were arbitrarily deducted for each error.) In the first recall he wrote the same paragraph at the rate of 19.8 words per minute, and at the beginning of the second recall at the rate of 10.8 words per minute. These measures taken at face value indicate that the experimenter, on the first day of the first recall, wrote the paragraph at a rate that took twenty-seven days practice to attain in the original experiment. In the second recall, which was undertaken twenty-five years after the first with no practice whatever in the meantime, and fifty years after the first learning, it took approximately eight days to secure the same skill. However, the two curves are similar throughout. Near the ends of the two recall curves it may be seen that the experimenter, in twenty-five to thirty days, was performing in a manner equivalent to that which had taken one hundred twenty-six days to accomplish in the original learning. At the end, both curves indicate that there was no limit to improvement in sight, and regardless of the age of the learner, improvement might go on indefinitely.

In copying the three hundred-word page of new material (Fig. 2), the story is somewhat different. In the original learning the experimenter copied the page at the rate of 4.4 words per minute. In the first recall he copied at the rate of 12.4 words per minute at the first practice, and in the second recall, on the first day, he wrote the page at the rate of 6.1 words per minute. Glancing at the two sets of practice curves it would seem that there is considerable similarity at the beginning. The experimenter was copying at a rate which, in three to seven days, took four times that many days to accomplish at the original learning. The first recall continues to show considerable improvement over the first learning, but the early promise of the second recall soon fades out and the curve

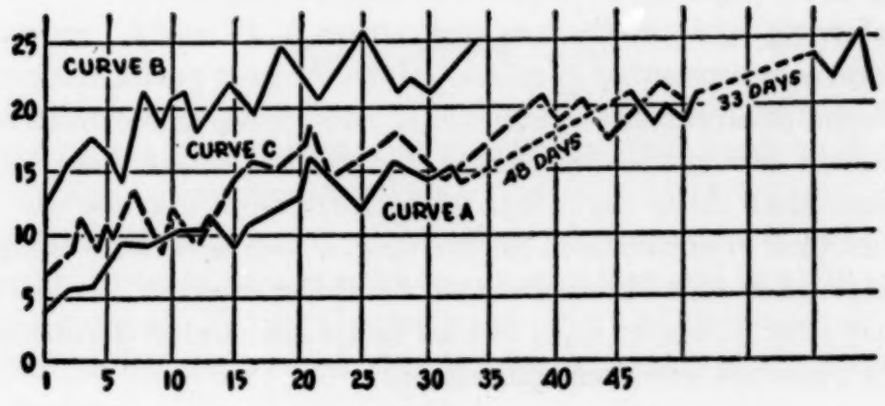

FIG. 2

almost parallels the original learning curve. Even this may be favorably regarded as significant considering the age of the experimenter. He believes, however, that the lack of more favorable progress is due to poor eyesight which makes reading from the printed page laborious and slow. This difficulty probably accounts for the leveling-off of the second recall curve, and probably indicates a limit to improvement in the rate of copying anything from the printed page where reading is involved.

Within the limits of this experiment certain observations may be made. The experimenter believes that age, as such, is not necessarily a deterrent to either new learning, or to recalling that which has formerly been learned. Physical handicaps do necessarily make a difference in performance. Muscular reflexes, where physical handicaps are not involved, perhaps may present a question for further investigation and study. This experimenter is inclined to believe that muscular reflexes in the aged, as in youth, show individual differences.

Some of the same factors that are present in learning at an earlier age are also present and effective at maturity. Interest, desire, and determination are potent factors. The will, as such, is a powerful stimulus. Competition against one's self is also very effective and may be more so in the aged than youth. While competition against rivals or competitors is a great stimulus to youth, the aged, as a rule, do not care who excels them.

In the case of the present experimenter, when he discovered that he could recall and even excel his previous record, and that he possessed the necessary skill to typewrite (for all practical purposes), he did not feel that extra exertion for more skill was worthwhile; although he does feel that he has by no means reached the limits of his capacity to improve in typewriting skill.

Referring again to the learning curves A, B, and C, we see ups and downs representing good and relatively poor performances, (as are found in all records of learning). The introspective notes which were kept during the experiment, along with casual observations, indicate that these variations in performances may be due to a current lack of enthusiasm for the task, a feeling of well-being—or the lack of it (sometimes referred to as the mood of the learner), or they may be due to some hidden factors in human performances which have not yet been discovered.

RECRUITING PROSPECTIVE RURAL TEACHERS AT THE MEN'S TEACHERS' COLLEGE TRIPOLI, LIBYA

IBRAHIM E. METAWEH[1]

UNESCO Mission, Tripoli

During the writer's successive contacts as a UNESCO technician with the Libyan education officials, the problem of effectively recruiting prospective rural teachers was raised. Teachers and students at the Teachers' College were included also in these contacts, and as a consequence, the idea of establishing an adequate instrument for recruitment was defined and clarified. The writer had noticed in the College several striking cases wherein students with serious defects had been admitted in the past, and since it seemed desirable to have some kind of device that would be effective in screening applicants, the idea of the interview test gradually took shape in the minds of the educational officials involved.

The purpose would be to raise the standards of prospective rural teachers admitted to the Teachers' College by means of a practical, objective instrument for screening applicants, namely, an interview test—one which might be standardized sufficiently to be used in other Teachers' Colleges in the country. The hoped-for result would be the selection of the fittest for the teaching profession from among the applicants, most of whom come from rural districts. The procedure adopted may be of interest to those working in similar regions and perhaps to others as well.

THE INTERVIEW CARD

Information, such as age, previous school grades, address, name of parents, etc., was obtained from every student applying for admission. Space was provided for each rater to mark the applicants on each of seventeen items. The items appearing on the card are presented below, supplemented by explanations and samples of the questions asked. The total number of points or marks arbitrarily assigned to each item and (for the main points) the number of

[1] Technician of Rural Education, UNESCO Mission, Tripoli.

questions under each are shown in parentheses. Students scoring between 40 and 49, were to be referred to a revision committee composed of the Dean of the College and two teachers in the Education Department. The minimum total to be obtained by every student was set at 50 marks out of a total of 100.

(1) *Physical Aspects (30 Points, 24 Questions)*

Failure in this item alone might result in ruling out the student. This would happen if he has an apparent physical defect sufficient to hinder him from expressing himself or from performing activities necessary to the teaching profession, or to make him marked among his students. This applies to the cases where the college cannot do anything at present to correct the defects. In such cases, however, the interview is continued and the student is referred to the revision committee. Elimination of cripples is necessary as we need the teacher to be a leader too, but this does not rule out the possibility of his being accepted in another profession.

(a) *General appearance (10).* Personal cleanliness and being well dressed without regard to the quality or price of clothing worn.

(b) *Speech and clarity (10).* Notice in the students' answers the fluency of speech, diction, ease in pronunciation, and form of sentences.

(c) *No tics and defects (10).* Notice here apparent lameness, dwarfishness, partial but apparent, or complete deafness, cross-eyedness, stuttering, stammering, lisping, awkward-walking, one-eyedness, facial deformities, and various tics, such as exaggerated gesturing, excessive movement of eye lashes, winking, excessive movement of head or shoulder and the scratching of various parts of the body.

(2) *Attitudes and Personality (25 Points, 28 Questions)*

(a) *Emotional balance and control (5).* (The student can be asked questions like the following, although it is not necessary to confine the questions to those mentioned here.) (*a*) What is your reaction if someone in a group does not like you? (*b*) What is your reaction if somebody differs from your opinion, and what are your feelings toward him? (*c*) How would you behave if you knew that one of your friends told others something about you that you did not like? (*d*) If you were late to a lecture and the teacher re-

fused to admit you, how would you behave? (*e*) What would you do if you were to fail this test?

(b) *Social interests (5).* (*a*) Do you like to be alone most of the time or do you get bored? Do you like to join groups? (*b*) Do you like to participate with others in organizing parties and public meetings? (*c*) Do you prefer to have a picnic alone or with others? To what extent? (*d*) What is the number of your acquaintances and friends? (*e*) Do you like to break the rules?

(c) *Leadership and guidance (5).* (*a*) If you are asked to participate in organizing a party, trip or social event, what rôle do you like to take? (*b*) If you like to accompany others on a picnic, do you like to have an opinion as to where you should go, or do you accept easily any place suggested by the others? (*c*) If you walk with others, do you like to be in front or in the middle, or do you not care about your order? (*d*) Have you studied the life of a leader? Who is he, and what did you admire about him? (*e*) If you are among a group of people who are discussing some affairs, what is your stand? Do you just listen to others? Do you participate in the discussion or do you guide the discussion?

(d) *Ability to comprehend and adjust (5).* (*a*) Complete the following sentence: "The relationship between the ox and the plough is like the relationship between the locomotive and ...." (*b*) Complete the following chain of numbers:

9–1–9–2–9–...–...–...–...–...
2–6–10–14–...–...–...–...–...
3–7–15–31–...–...–...–...–...

(*c*) Ali is Ibrahim's cousin on both sides (mother and father). What is the relationship between the father of Ali and the mother of Ibrahim? (*d*) The speed of Ali is half the speed of Ahmed and Ahmed's speed is one and a half as much as the speed of Elsayed. What is Ali's speed if Elsayed's speed is 6 kilometres per hour? (*e*) What would you do if the neighbouring house were on fire? (*f*) What would you do if you were asked to talk on a subject with which you were not acquainted? (*g*) Supposing that you were late for an important appointment where your presence was necessary, how would you act? (*h*) If you were given fifty Libyan pounds, what would you do?

(e) *Independence and self-reliance (5).* (*a*) Do you seek the help of others regarding your personal problems? (*b*) If you find it difficult to understand a lesson or an arithmetic problem, what

do you do? (*c*) Do you like to have a tutor to help you during the scholastic year, or near the examination time? (*d*) Do you like to choose your clothes and their colors personally, or do you accept what the family buys for you? (*e*) If you are to go to a place which you have not visited before, do you like to have somebody you know to accompany you, or how would you act?

(3) *Interests and Attitudes Towards Teaching (25 Points, 31 Questions)*

(a) *General interest in teaching (5).* (*a*) Do you like being a teacher? (*b*) What makes you like teaching? (*c*) Do you ever help younger brothers, sisters or relatives with their lessons? How much time do you spend on this help? (*d*) Do you enjoy helping them, or are you bored with this task? (*e*) Who is your favourite teacher, and why?

(b) *Interest in working with children (5).* (*a*) Do you participate with children in their games and play? (*b*) What are your feelings while playing with children? (*c*) Do you feel that children like you? (*d*) Do you find pleasure in talking with children or are you annoyed by it? (*e*) Do you like to listen to children's stories? (*f*) Do you enjoy telling stories to children and do children listen to you? (*g*) Do you enjoy taking your younger brothers, sisters or relatives on a picnic, to a movie, or to visit relatives? (*h*) If you are at home while children are playing, how do you feel about it? Are you annoyed with their play and noise? Do you like to participate? Do you enjoy watching them?

(c) *Interest in physical training (5).* (*a*) What games do you play? (The student should be asked about the way in which he plays the games that he knows and about their rules.) (*b*) What games do you prefer and why? (*c*) How frequently do you play these games? (*d*) Do you like watching games? What kind of games do you like to watch? (*e*) Do you prefer individual, group, or mass games?

(d) *Interest in reading and growth (5).* (*a*) What do you prefer to read apart from your assigned school books? What kind (religious, scientific, social . . .)? (*b*) What books did you read last year? (*c*) What benefit did you obtain from these books? (*d*) Do you read newspapers and, if so, which? What sections of the newspapers do you prefer to read? (*e*) Do you read magazines and, if so, which? What sections do you prefer to read?

(e) *Extent of participation in clubs and school organizations (5).* (*a*) Are you a member of any school organizations and, if so, which? What are your activities in these organizations? (*b*) If you are not a member, why not? (*c*) Are you a member of an organization outside the school, in the village or town? What is your rôle in it? (*d*) Are you a member of a club and, if so, which one? (*e*) How much time do you spend in the club? (*f*) What activities do you carry on in the club? (*g*) If you have not joined any club, why not? (*h*) What benefit do you derive from joining different organizations?

(4) *General knowledge (Historical, Hygienic, Economic, Political; 20 Points, 57 Questions)*

(a) *Knowledge about the locality (5).* When is your birthday? In what place were you born? What is the population of your birthplace? What "markaz" and what Province does your birthplace follow? Is your Province famous in producing certain goods? Your home town's celebrity? Name of Wali or Omda? Hospitals and schools in the area? Transportation routes from your home town to different places? Population of the Province? People's vocations?

(b) *Knowledge about Libya (5).* The Libyan revolutions for independence? Who led each? When did they take place? What did these revolutions achieve for Libya? The date of independence? Number of deputies in Parliament? Number of Senators? Who are the Prime Minister and Ministers? Population of Libya? Largest cities, their population and celebrities? Names of some villages? Names of some industrial cities? Routes of transportation in Libya?

(c) *Knowledge regarding the Arab world (5).* What is the Arab League? Member states? Capitals of each? Population of each? Means of transportation to each? Kind of government in each? Headquarters of the Arab League? Palestinian problem? Famous men and products of each Arab state? Information about Morrocco and Tunisia, Algeria and Imaret Alkwait? Why is it that the third and fourth states named are not members in the Arab League?

(d) *Knowledge about the world (5).* (The student should be asked about the last World War and its effects.) The states which were involved in this war? The capitals of these states? Important

figures about these states? The relationship between Libya and these states? Important products of these states? Their famous men? Kind of government? The United Nations, Specialized Agencies and Security Council? The functions of each?

PROCEDURE AND APPLICATION OF THE TEST

Each interviewee was rated by a committee of at least three persons from the staff of the Teachers' College, one with a scientific background, the second with an educational or social-studies background, and the third with a linguistic or artistic background.

Directions to the committee members caution them not to begin to question a student until they establish rapport and are sure that he is at ease. Students can be assured of permissiveness by talking to them about simple and ordinary things in a friendly way. It is strictly forbidden to scold a student or to be sarcastic to him during the interview, no matter what his answers are. He is not to be hurried in giving his answers or interrupted in his train of thought. The committee members are reminded that the students examined are young and in the process of growth, and that the conditions of their environment, their way of life, and economic status should be taken into consideration while examining them. It is not advisable to ask the students about peculiar features of areas that they have not visited. To give them adequate time to reveal their interests and attitudes, and at the same time to guard against exhausting them, the period of the interview was held to ten or fifteen minutes.

Members of the testing committees are free to start the test without adhering to the consecutive order of questions given. They might like to start by inquiring of the student about his locality. Also, it is not compulsory to follow the patterns of questions and directions stated. They are given as examples but the test permits flexibility.

When the test interview with its four main points, one hundred marks, and one hundred and forty questions was finally agreed upon, the writer presented a copy with the instructions to the appropriate authorities. A committee of four that was formed to examine it approved it without modifications. Fifty copies were mimeographed, twenty-seven of which were distributed among the teachers who conducted the test. Eight examining committees were

formed at the Teachers' College made up of three teachers each, all of whom studied the test in advance.

RETURNS AND EVALUATIONS

Five hundred and thirty-six students applied for admission varying in age from eleven to twenty-four, approximately three out of every four students coming from rural areas. The average age of the students who applied was sixteen years and ten months. The average age of the admitted students was sixteen years and nine and a half months. Twenty-six students were not admitted because they were thought to be too young or too old. (Table I.) The age of seventeen years was thought to be most suitable for prospective teachers as they will have to stay in college for four years, and by the time they graduate at the age of twenty-one, it is highly probable that they will be mature enough for teaching.

Of the five hundred and ten applicants not disqualified because of age, 60.4 were admitted (57.5% of the total number). It may be that some of the more than one-third who were disqualified might have become successful in college and as teachers. This will never be known, and only time will tell how many of those who were admitted will be successful. Of those disqualified because of a score of less than 50, some extreme cases were—ten students who were extremely short, four one-eyed, seven stuttering and stammering

TABLE I.—AGE OF APPLICANTS ADMITTED AND NOT ADMITTED

Age	Applicants	Admitted	Not Admitted
12 years or less	10	0	10
13 years	9	5	4
14 years	9	5	4
15 years	34	21	13
16 years	149	91	58
17 years	181	108	73
18 years	93	58	35
19 years	25	13	12
20 years	6	4	2
21 years	4	3	1
22 years and over	16	0	16
Total	536	308	228

cases, five lame, and twenty apparently suffering from malnutrition.

This project is radical in recruiting intended teachers in this region where admission procedures have long remained unorganized. There were some doubts as to the necessity for this new interview test among some officials and teachers since they were well enough satisfied with conditions as they were. Throughout the test and after it was tried out their attitude changed. The writer is glad to report that they are now enthusiastic about it and are willing to participate in every phase of it in the coming years. This in a way proved the effectiveness of the test.

Since it had not been customary to obtain the exact birth dates of applicants nor a transcript of credits from the schools they had previously attended, it is as yet impossible to correlate ages and marks with the results from the test interview, but it is proposed to correct for this deficiency in the future. Similarly to test the instrument's validity correlations will be obtained with the results of the diploma examinations and with marks in practice teaching. Certain revisions will be made in the interview test to include more questions, particularly in the field of artistic interests, which were not covered. Work is soon to start on developing an intelligence test, the results of which will be used to supplement the other information used as criteria for admission.

Lack of social interest, independence, and self-reliance were apparent from the interview test results coupled with the observations of the testing committees. The reasons for this are difficult to ascertain, but it is felt that the lack of extra-curricular clubs might, in part, be responsible. So far there are five newly formed clubs in the College: the Historical Club, the Educational Club, the Journalism and Magazine Club, the Athletic Club, and the Scientific and Cinema organization. Although they are cultural in nature, social attitudes, independence, and self-reliance are involved.

The first use of an instrument in this country has been most gratifying since it was a coöperative enterprise that seemed to meet the needs of the situation as we knew it. As a result of its use and of other projects under way, we can look forward to a gradual upgrading of the teachers in Lybia, and hence, of its educational system.

THE INTELLIGENCE TEST IN SCHOOL USE: SOME PERSISTENT ISSUES[1]

J. B. STROUD

State University of Iowa

The first successful intelligence test was designed with little regard to definition or theory. Binet proceeded in an empirical fashion to ascertain kinds of tests that would discriminate among criterion groups. Kuhlmann, Terman and most of the early workers adopted the Binet procedure. On the other hand we are all familiar with various analytical approaches to the problem of intelligence test design. Thorndike's *area, altitude* and *speed*; Spearman's *g* and *s* factors; and Thurstone's *primary mental abilities* are cases in point. Thus, Spearman wrote, "When a knowledge of *g* and *s* factors makes it possible to construct tests which measure *g* and *s* at the maximum, an accurate measure for *G* for any individual can be obtained." This, he maintained, is far better than an average derived from a miscellany of unrelated traits (*7*).

Thus, the history of the intelligence test reflects two major approaches in test design: (1) the empirical and (2) the analytical. In the former, the psychologist attempts to maximize differences among criterion groups. This is largely a utilitarian approach. In the analytical approach the psychologist investigates the nature of intelligence, with little immediate concern for the practical. Obviously it was Spearman's hope that the two approaches would ultimately amount to the same thing. As yet this hope has not been fulfilled. Test designers have generally resorted to the Binet method of maximizing differences among criterion groups. The advancement in design is reflected largely in statistical and logical refinement of Binet's general plan. Perhaps the most notable departure is the PMA battery. Here one gains the impression, from the somewhat limited research, that insofar as the tests of the PMA battery are factorially pure, most of them assess mental abilities not closely correlated with school achievement (*11*).

Binet regarded the MA as a general average. Perhaps the general average is still the most significant feature of the major tests.

---

[1] Presidential Address, Division of Educational Psychology, 1956.

However, most general tests have featured batteries of subtests permitting part scores and resulting profiles. This practice tends to insure systematic sampling of intellectual abilities and thus to improve the "general average." Ostensibly the purpose of this feature is to provide for differential diagnosis and prediction. Although discussion of the analytical features of intelligence scores has sharpened within the last decade or so, the usefulness of the part scores in school situations still suffers from a lack of detailed research on the differential power of the constituent subtests.

Nonverbal tests came into the picture originally as a means of examining persons presenting language handicaps. Most of the new general intelligence tests within recent years have featured both verbal and nonverbal scores. Wechsler's recent and impressive Intelligence Scale for Children, affectionately referred to as the WISC, and the more recent and equally impressive Lorge-Thorndike tests are cases in point. Insofar as we know, at least insofar as I know, this trend in test design is dictated more by fashion than by empirical evidence attesting its usefulness in general school situations. I suspect one would have some difficulty marketing group intelligence tests today that did not contain nonverbal features. I also am inclined to suspect that their worth as nonverbal tests in school practice has not been demonstrated by research. What the clinical psychologist is able to make of the results of this kind may be another matter.

Most of the earlier nonverbal tests were seriously lacking in validity as intelligence tests. We have just completed a correlation analysis of the WISC on about eight hundred school children, grades three to six inclusive. In our results the nonverbal tests predicted academic achievement quite as well as the verbal tests. The attested usefulness in school situations, however, of the Wechsler nonverbal tests may well stem from the fact that they are valid tests of intelligence, rather than from their nonverbal character. Quite conceivably additional verbal tests of comparable quality might be substituted for the nonverbal ones without detriment to the battery.

Our analysis failed to yield much in the way of differential prediction of the achievement measures used, reading, arithmetic, and spelling. It is true that four or five of the eleven subtests gave better prediction than did the others of the three achievement measures. But this apparently resulted from the fact that they

were better tests rather than from qualitative differences. A multiple regression analysis was performed based upon within-grades intercorrelations between the eleven WISC subtests and the three achievement measures. *Beta* weights were tested for statistical significance and least significant variables successively eliminated. It was found that four WISC subtests—Arithmetic and Vocabulary among the verbal tests, and Block Design and Object Assembly among nonverbal tests—accounted for most of the correlation between the WISC and the achievement measures. The within-grades least squares multiple correlation between all eleven WISC tests and reading, for example, was 0.74. The within-grades least squares multiple correlation between these four subtests, Vocabulary, Arithmetic, Block Design and Object Assembly, and reading was 0.73. The correlation between the unweighted raw score sums of the four subtests just referred to and Reading was 0.69. Comparable results were obtained for arithmetic achievement and for spelling achievement. It is a particularly striking fact that for all practical purposes the same four WISC subtests are the most effective predictors of each of the three achievement tests. For cross validation purposes the same kinds of analyses were carried out on a sample different from that on which the regression equations were derived, but drawn in the same way. The results obtained with the independent sample were strikingly similar to those obtained on the original sample (*12*).

Incidentally, analyses of this kind seem to be indicated for differential aptitude tests. It has been suggested that in view of the high correlations obtained between general aptitude tests and intelligence scores and also between general achievement tests and intelligence scores, the intelligence test is no longer required in school practice insofar as prediction of achievement is concerned. It is true that prior achievement predicts future achievement at least as well as does the general intelligence test. But it is equally true that the intelligence test which can be administered in a fraction of the time and at a fraction of the cost predicts future achievement almost as well. The merits of the argument must rest pretty largely upon the nature of the evidence of differential prediction. This brings us back to the intelligence score. How much have we improved upon Binet's "general average," beyond obtaining better averages? The value of pat scores on a test—verbal, nonverbal or others, depending upon the analytical properties of the

test—must be assessed in terms of differential prediction, actually substantiated by research. Perhaps it is true that the general achievement test and the general intelligence test each tells us something unique about the pupil and that arguments about one's replacing the other could well be forgotten. However, and for the time being, it may be well to bear in mind that the intercorrelations among achievement scores in various school subjects are relatively high. In view of this fact it will be difficult to do a great deal with differential prediction. Consider a medical aptitude and a legal aptitude test, for example. One has a medical aura, the other a legal one. But it still may be true that they could be pretty well interchanged so far as their basic purposes are concerned.

Another issue at least related to the intelligence score is the generality of the mental age. The mental age score, whether determined by Stanford-Binet method or that derivable from Wechsler tests by multiplying the obtained IQ by CA, is, in effect, an average score. Are some MA's, say, of ten, better than others? Do MA's of a given magnitude vary qualitatively? In our work we have sought to determine whether or not MA's of a given magnitude vary qualitatively with the brightness of the pupils achieving them. A related but somewhat different question is whether or not we can make the same predictions relative to current school achievement of bright and dull pupils earning like MA's. Of course, we would predict much greater future achievement for the younger pupils. In one investigation Kolstoe (*8*) formed matched pairs of bright and dull pupils achieving equal, estimated true Stanford-Binet MA's. The mean Stanford-Binet MA was 11.3. The mean CA of the dull members was 15.5, giving a mean IQ of 79. The mean CA of the bright members was 9.3, giving a mean IQ of 126. The bright and dull pupils in the sample thus formed were then compared on the basis of the eleven WISC subtests and the PMA subtests. Except for PMA subtest Number, differences between the bright and dull pupils on the seventeen subtests were small and generally not significant. The dull, older pupils were far better in performance on PMA subtest Number.

In a companion investigation, Bliesmer (*2*) compared bright and dull pupils, matched in MA, on reading achievement. On reading subtests Recognition of Main Ideas, Drawing Inferences, and Total Comprehension, the bright, younger pupils excelled by significant amounts.

The school placement of the older pupils in this sample was three or more grades higher than that of the younger pupils. Owing to the hierarchical character of the arithmetic curriculum, the younger pupils had had no instruction in some of the operations which the older pupils had studied. This may account for some or all of the superiority of the older pupils in Number. One might think, *a priori,* that young, bright pupils would surpass (in achievement) older, dull pupils of comparable MA's. They are likely to be better motivated, they have been expected to do well in school, they are more likely to be accustomed to success. In general they have had better early home training and have had more encouragement and stimulation at home. Of course, these advantages may be offset by the fact that at the age level at which the work just mentioned was done the older pupils had spent twice as many years in school as the younger ones.

At the present time we are witnessing a growing interest in the education of gifted pupils. This is obliged to sharpen interest in the nature of the mental abilities of the gifted, especially the highly gifted. Perhaps we would be inclined to infer from the results of the work on the generality of the Mental Age that ten-year-old pupils earning MA's of fifteen would behave on such tests as the WISC and the PMA much as fifteen-year-old pupils of average intelligence.

The work on the mental age concept just referred to certainly suggests that for the age ranges involved, the MA as a unit of measurement possesses a high degree of generality. This is not to defend the Stanford-Binet procedure of designing intelligence tests. I do mean to defend the MA as a highly useful unit of measurement. The Wechsler method of designing intelligence tests, while it may not produce a superior result, appears to be more efficient and more economical.

The logic of the argument that the MA as a unit of measurement suffers from known or suspected inequalities of its units is not impressive. The meaning and usefulness of Height Age or Weight Age or Dental Age are unaffected by the fact that these units prove to be unequal when tested by a metrical scale. A height age of nine is scarcely more meaningful in school work because we know how much this is in metrical units. A Mental Age, or for that matter a Reading Age, of nine is not less meaningful because we

are unable to say how much intelligence or how much reading ability it is.

Moreover, the inequality of successive MA units, or our lack of knowledge regarding their equality, argues nothing with respect to the equality of IQ's obtained at successive age levels. The IQ defines a pupil's mental development in relation to the mental development of other pupils of like age in the general population. In a sense it defines the rate of mental growth. We may say that a six-year-old pupil who achieves an IQ of 125 develops at the rate of 1.25 of a mental year for each calendar year. At this age his attainment equals that of the average pupil seven and a half years old. He has gained a year and a half in six years. He would be expected to gain another year by the time he is 10. We could also say he is approximately 1.5 S.D.'s above the mean for his age, or that he exceeds about 93 per cent of the pupils of his age in the general population. The fact that the MA score from which the IQ is derived at successive ages may not be equal has nothing to do with a pupil's maintaining his relative position in his age group.

Terman's concept of validity of the IQ implies an absence of systematic variations in IQ with age. Given constant means and S.D.'s over an age range, a given IQ must signify the same degree of brightness at successive age levels within this range. This assumes nothing with respect to the equality of the MA units from which the IQ's are derived. Unfortunately, test designers have not always been as careful as they should have been to publish S.D.'s for successive ages over the range of years covered. In the absence of this or comparable information we cannot tell whether or not an IQ of a given magnitude at one age level is strictly comparable numerically to one of like magnitude at another age level.

One of the more troublesome practical problems is the comparability of IQ's obtained on different tests—one might say needlessly troublesome. It should not have been a difficult matter at all for test designers to have provided school psychologists or other school officials with the variability data necessary to the establishing of numerically equivalent scores on various tests.

In school work there constantly arises the question of the significance of fluctuations in IQ from one testing to another, no matter whether the same or different tests are used on different occasions. When test designers, as Terman and Wechsler have done, provide the necessary data, the school psychologist can read-

ily acquaint teachers with the amount of fluctuation required on a given test to become significant at given levels of confidence. There is no excuse for test authors' not doing this.

Traditionally our schools have been interested in several kinds of growth data. The variables for which we have the most satisfactory measures are mental growth, educational growth, and physical growth. It is of interest to note that these three types of growth possess considerable stability. Correlations between intelligence scores earned at eight and at eighteen years of age run around 0.85, as seen, for example, in Bayley's work (*1*). Correlations between measures of stature, obtained at eight and at eighteen years are of somewhat comparable magnitude (*9*), as are those between measures of educational achievement taken ten years apart.

We also know from an abundance of research that intelligence scores and educational achievement scores are rather closely intercorrelated, and that neither of these is closely intercorrelated with measures of physical growth. Moreover, we can improve prediction of educational achievement somewhat by combining intelligence and past achievement scores. We do not, as the work of Gates (*6*) showed more than twenty-five years ago, improve prediction by combining intelligence scores with various measures of physical growth.

However, the statistical method ordinarily used, that of multiple correlation, precludes our obtaining a lower correlation than that obtained between intelligence alone and achievement. If, on the other hand, we form simple averages involving MA's and various physical growth age scores, such as Height Age, Weight Age, Grip Age, and so on, and correlate the averages thus obtained with achievement, we do stand to obtain a much lower correlation. Indeed, if as many as five or six physical age measures are thus combined with MA and the resulting averages correlated with achievement, the coefficient should approach in magnitude the coefficient obtained between physical growth measures and achievement.

This is not to controvert the importance of physical growth in the conduct of schools. Physical growth features in the health and well-being of children. It is probably an important factor in their social and emotional adjustment. Certainly it would appear that any serious deviations should pose personal problems to which

adjustment must be made. But we do not improve, but lower, prediction of achievement by combining physical age scores with MA scores.

For about as long as I have been acquainted with the subject at hand, there has been some expressed concern about the effect of reading proficiency on group intelligence scores. To be sure this is primarily a practical problem. But I suspect it could be argued that some of us have published papers on problems that had less practical—or theoretical—significance. In a way it is a bit strange that more work has not been done on this problem, if for no other reason than that of allaying the concern of teachers and school officials.

Some work of one of our own students on an MA thesis suggests that a slow reader who has good reading comprehension is penalized, so to speak, on timed group tests (*10*). No evidence was obtained to suggest that the pupils scoring low on reading comprehension, whether fast or slow in rate, were penalized in any way. It seems probable that a facile reader who enjoys excellent reading comprehension would outperform on timed group intelligence tests, an equally intelligent pupil who reads slowly. Whether such an outcome would help or hinder prediction would depend upon the relationship between reading rate and educational achievement. One would suspect that the relationship is far from close. It is conceivable, as Durrell (*4*) pointed out long ago, that since reading proficiency is correlated with achievement, the presence of a reading factor in group intelligence tests may actually improve their predictive power.

Obviously, we would not expect the effect of reading proficiency on group intelligence test scores to be the same for all such tests. We would anticipate that the obtained effect would, among other things, prove to be a function of the speed-power variable.

By way of closing, some brief remarks are made apropos of the social status question in intelligence testing. Without doubt, the publicity given this question by the work of Eells, Davis and colleagues (*5*) and the earlier work of Cattell (*3*) should have a wholesome effect upon test construction. As a group, test authors in the past were probably guilty of some degree of carelessness with respect to this problem. It is perhaps too early to attempt to resolve the important issues raised. The need is urgent to understand the meaning of social class differences in test scores as

psychological phenomena. The use of such expressions as "culture fair" and "culture bias" gives the impression that these differences are the unwitting work of test authors. I say unwitting because surely no one would have done this by design. Obviously, if these differences are chargeable to test authors, they are unimportant as psychological phenomena, save as they become something to be studiously avoided. On the other hand, we may find that the problem is really too big for the test author. It may turn out that the kinds of cultural impact associated with social class differences affect the course of *mental development* of children as well as their performance on intelligence tests.

Social class differences are real differences—substantial psychological phenomena with which schools and society must deal. It may be that we cannot build valid intelligence tests which will not at the same time discriminate among the social classes. Or, if we start the other way round by designing tests which will not discriminate among social classes, we may find that the tests are poor predictors of academic achievement.

REFERENCES

1) N. Bayley, "Some Increasing Parent-Child Similarities During the Growth of Children," *Journal of Educational Psychology,* 45, 1954.

2) E. P. Bliesmer, "A Comparison of Bright and Dull Children of Comparable Mental Ages with Respect to Various Reading Abilities," *Journal of Educational Psychology,* 45: 321–31, 1954.

3) Raymond B. Cattell, "Classical and Standard Score IQ Standardization of the I.P.A.I. Culture-Free Intelligence Scale 2," *Journal of Consulting Psychology,* 15: 154–59, 1951.

4) D. D. Durrell, "Influence of Reading on Intelligence Measures," *Journal of Educational Psychology,* 34: 165 ff., 1943.

5) Kenneth Eells, et al. *Intelligence and Cultural Differences, Chicago,* University of Chicago Press, 1951.

6) A. I. Gates, "The Nature and Educational Significance of Physical Status and of Mental, Physiological, Social and Emotional Maturity," *Journal of Educational Psychology,* 15: 329–58, 1924.

7) Katharine B. Greene, "General Intelligence and Its Measurement," *Review of Educational Research,* 2: 277, 1932.

8) O. P. Kolstoe, "A Comparison of Mental Abilities of Bright and Dull Children of Comparable Mental Ages," *Journal of Educational Psychology,* 45: 161–168, 1954.

9) H. V. Meredith, "The Prediction of Stature of North European Males through the Elementary School Years," *Human Biology,* 8: 279–83, 1936.

10) J. F. Rummel, "The Effect of Reading upon Group Intelligence Test Scores," Master's Thesis, University of Iowa, 1947.

11) D. C. Shaw, "A Study of the Relationships between Thurstone Primary Mental Abilities and High School Achievement," *Journal of Educational Psychology,* 40: 239–49, 1949.

12) J. B. Stroud, Paul Blommers, and Margaret Lauber, "Correlation Analysis of WISC and Achievement Tests," *Journal of Educational Psychology,* 48: 18–26, 1957.

THE INFLUENCE OF INTENTION ON THE RETENTION OF SCHOOL MATERIALS

DAVID P. AUSUBEL, SEYMOUR H. SCHPOONT, and LILLIAN CUKIER

Bureau of Educational Research

University of Illinois

Extrapolating from general principles of educational psychology and from related experimental evidence, many educators and psychologists implicitly believe that intention to remember facilitates the retention of learned verbal materials. Actually all of the evidence bearing on this issue is ambiguous and inconclusive for one or both of two reasons: the experimental arrangements have been such as (1) to induce intentions to learn rather than to retain, or (2) to make impossible the isolation of the effects of intention on learning from its effects on retention. The present study was designed to overcome these interpretive ambiguities by both introducing *explicit* intentions to remember for an extended period of time and by experimentally isolating the effects of such intentions from original learning. The latter result could only be insured by introducing intentions to remember *after* completion of the learning task.

Numerous studies have shown that intention to remember influences the quality and longevity of retention. When subjects learn material with the expectation of recalling it for a designated period of time, recall is superior for the expected as against either a shorter or longer interval (*2, 3, 8*). Lester demonstrated that retention is enhanced by expectation of recall and by foreknowledge of the occurrence and possible effect of interpolated materials (*5*). Unfortunately, however, since differential intentions to remember were introduced at the time of original learning, these experiments did not isolate the effects of intention on the quality or quantity of what was learned in the first place from its effects on retention per se.

Another group of studies (e.g., *6, 7*) has shown that retention is superior and retroactive inhibition less marked when practice is accompanied by "intent to learn" than when learning takes place

incidentally. That this difference is largely attributable to superior original learning, however, is demonstrated by the fact that it no longer prevails when Ss are equated for original mastery of the learning task (*1*). Heyer and O'Kelly (*4*), on the other hand, showed that learning under conditions of strong ego-involvement did significantly increase retention even though their ego-involved and control groups could not be differentiated on the basis of original learning outcomes. It is quite possible, however, that the superiority of the ego-involved group on original learning was masked by the inhibitory effect of intense ego-involvement on initial test performance. In any event, the evidence yielded by all of these studies can at best be indirect, because in each instance an explicit experimental set was induced to *learn* material for immediate reproduction rather than to *retain* it for an extended period of time. Furthermore, even if ego-involved intention to learn leads to no demonstrable quantitative superiority in original learning, it may still affect retention by influencing certain qualitative aspects of the learning process.

METHOD

Standardization of learning material and tests of retention. The learning material used in this study was a passage of approximately one thousand four hundred words on the history of opiate addiction. It was selected on the basis of being generally unfamiliar and inaccessible to undergraduate students, but yet learnable and relatively interesting. Two equivalent multiple choice tests on this material were constructed as follows: fifty-six five-choice test items were administered to a group of fifty-five undergraduate students in education three days after they had studied the passage for twenty-five minutes. Following item analysis, the order of items was rearranged so as to match odd and even items with respect to difficulty and type of question. The revised form of the test was then administered to a different group of twenty-nine undergraduate students in education. The range of scores (number correct) on this test was from thirteen to forty-five with a mean of thirty-three and a standard deviation of 7.8. The uncorrected odd-even coefficient of equivalence was 0.76. For the main experiment the odd and even items respectively constituted the two equivalent forms.

Experimental procedure. Our experimental population consisted

of ninety-seven undergraduate students enrolled in four sections of an educational psychology course at the University of Illinois. Of these, fifty-three were in the control group and forty-four in the experimental group. In order to minimize the possibility and consequences of communication between groups, the control group was handled during the Fall semester and the experimental group during the following Spring semester. The experiment was rationalized as a laboratory exercise and was performed during scheduled class hours. The control Ss were first instructed to read the passage at their normal reading speed (approximately eight minutes), and then to utilize the remainder of the twenty-five available minutes to commit the facts and concepts in the passage to memory. It was explained that a multiple choice examination on the material would be given immediately after the reading was completed. Ss were assured that the results on the examination would not be made public and would not affect their grades in the course. They were promised, however, that their scores would be returned to them individually, and that the median and frequency distribution of the class scores would be placed on the blackboard so that they could compare their own scores with the class norms. Fourteen days after the original learning and test, an unannounced (surprise) retest was given, utilizing the equivalent form of the test. Hence the control group had no explicit intention to remember the material for the second test.

During the second semester the same procedure, with one exception, was followed for the experimental group. Immediately *after* the Ss finished the first test, they were informed that a second test on the same material would be given in fourteen days. This announcement presumably induced explicit intentions to retain the previously learned material for an additional two weeks without in any way affecting the quality or quantity of original learning.

RESULTS

Means and standard deviations of learning (test) and retention (retest) scores for control and experimental groups are shown in Table I. Comparison of test and retest scores indicates that a significant degree of forgetting ($p < 0.001$) occurred in each group during the two week interval. The mean learning and retention scores made by the two groups were also significantly greater in all

TABLE I.—COMPARISON OF CONTROL AND EXPERIMENTAL GROUPS ON LEARNING AND RETENTION TESTS

	Control Group		Experimental Group	
	Test	Retest	Test	Retest
Mean	20.1	13.2	19.5	13.6
S.D.	2.53	3.06	2.56	2.58
Corrected mean*	13.4	6.5	12.8	6.9
Percentage retained†	48.5		53.9	

* The corrected means were computed by subtracting 6.7 (the mean score of subjects who took the test without studying the material) from the group means of raw scores (number of correctly answered items).

† Corrected means were utilized in computing the percentage of retention from test to retest.

instances than the test scores made by different Ss who answered the test questions without prior study of the learning passage. The product-moment correlation between test and retest scores was 0.45 ($p < 0.01$).

Since the means, standard deviations, and variances of the two groups are not significantly different from each other on the learning test, we are justified in comparing control and experimental groups with respect to percentage of retention from test to retest. The respective percentages of 48.5 and 53.9 show only a negligible difference in degree of retention which does not even begin to approach statistical significance.

DISCUSSION

The results of the present study do not support the widely held view that intention to remember facilitates the retention of learned verbal material. In reconciling these findings with previously reported studies in this area (*2, 3, 5, 8*), we can only surmise that intention to remember in these latter studies did not influence the retention process per se, but affected various quantitative and qualitative aspects of original learning which indirectly increased the availability of the learned material at a later date.

Obviously, if a student learns material more thoroughly in the first place because he expects to be tested on it several weeks later he will have more available to remember than a student who originally learns less diligently because he does not expect to be

retested. However, if the intention to remember is introduced *after* learning takes place it apparently has no influence on the longevity of the learned material. Intention to remember, in other words, affects retention by enhancing learning rather than by increasing the stability of existing memory traces; or, to put it negatively, intention *not* to remember material beyond the date of the examination leads to little subsequent retention not because of any direct influence of this intention on memory traces but because such intentions are invariably associated with superficial and cursory methods of original learning. Hence, teachers who wish to increase the longevity of learning by maximizing intention to remember can only do so by inducing the appropriate intention at the time of original learning. Admonitions to remember once learning is completed, even if they lead to corresponding intentions in the student, do not alter the subsequent fate of memory traces.

One possible qualification of these findings should be considered. Intention to remember in our Ss might have been relatively weak in the absence of tangible material rewards for superior retention. High retention scores, after all, did not "count" in the sense of improving students' grades. On the other hand, level of ego-involvement was maintained at a reasonably high level by intrinsic desire to coöperate in a research study relevant to the subject matter of the course and by Ss' incentive of scoring favorably in relation to class norms.

SUMMARY AND CONCLUSIONS

An attempt was made to ascertain whether explicit intention to remember facilitates retention of learned school materials apart from its influence on original learning. An experimental group of undergraduates studied a historical selection for twenty-five minutes and was tested on same immediately afterwards by means of a multiple choice test. *After* this test, an explicit intention to remember was induced by announcing that an equivalent form of the test would be given two weeks later. The same procedure was followed with a control group except that an unannounced retest was given.

The two groups were not significantly different in mean learning scores or in the percentage of material retained from test to retest. It was concluded that intention to remember facilitates retention by enhancing original learning rather than by increasing the stability of existing memory traces.

REFERENCES

1) W. C. Biel, and R. C. Force, "Retention of Nonsense Syllables in Intentional and Incidental Learning," *Journal of Experimental Psychology*, 32: 52–63, 1943.

2) F. D. Boswell, and W. S. Foster, "On Memorizing with the Intention Permanently to Retain," *American Journal of Psychology*, 27: 420–426, 1916.

3) M. T. Geyer, "Influence of Changing the Expected Time of Recall," *Journal of Experimental Psychology*, 13: 290–292, 1930.

4) A. W. Heyer, Jr. and L. I. O'Kelly, "Studies in Motivation and Retention. II. Retention of Nonsense Syllables Learned Under Different Degrees of Motivation," *Journal of Psychology*, 27: 143–152, 1949.

5) O. P. Lester, "Mental Set in Relation to Retroactive Inhibition," *Journal of Experimental Psychology*, 15: 681–699, 1932.

6) J. Peterson, "The Effect of Attitude on Immediate and Delayed Recall: a Class Experiment," *Journal of Educational Psychology*, 7: 523–532, 1916.

7) W. C. H. Prentice, "Retroactive Inhibition and the Motivation of Learning," *American Journal of Psychology*, 56: 283–292, 1943.

8) M. N. Thisted, and H. H. Remmers, "The Effect of Temporal Set on Learning," *Journal of Applied Psychology*, 16: 257–268, 1932.

THE FACTORIAL INVARIANCE OF PURE-FACTOR TESTS[1]

JOHN W. FRENCH

Educational Testing Service

Current practice in assembling batteries of pure-factor tests has been to select, for each factor, one or two tests that have been found by factor analysis to have high loadings on that factor and low loadings on all others. Factorial invariance is assumed. That is, if several tests were to be selected to represent each of a number of factors and tried out with a new group of subjects, it would usually be assumed that the factors would reappear clearly defined by the expected tests. The analysis described in this paper was carried out to check on the invariance of a set of pure-factor tests selected for possible use in making educational predictions.

The tests making up the experimental battery of pure-factor tests were suggested by tests used in the studies covered in a survey of the factor-analysis literature (*2*). In general, three tests were selected as the purest and most reliable measures of each of thirteen of the best-established aptitude and general achievement factors.

Deviating from this pattern, Reading Comprehension and Mathematics Reasoning were used as measures of two factors each: Verbal Comprehension and General Reasoning for the former, Number and General Reasoning for the latter. Although these two are not pure-factor tests, they were included because of their known validity for a number of purposes. The Carefulness factor represents another deviation from the pattern of three tests for each factor. Since not much confidence was felt in the way that the scores for the Carefulness factor would turn out, five tests were tried. Four are used in the analysis; the other was eliminated, since eighty-five per cent of the subjects received a zero score. It should be pointed out that the Carefulness scores were derived from the tests of other factors. However, it was believed that there would be little distortion of the factors produced by experimental

[1] Grateful acknowledgment is made to the staff of the U. S. Military Academy for their coöperation in scheduling and administering the tests and in supplying data on course grades. Without this kind of coöperation the study would not have been possible.

dependence, because the four Carefulness scores bore no consistent relationship to the scores on the four tests from which they were derived.

THE TESTS

The tests are described in the order in which they were administered. The first four were a part of the entrance tests given five months earlier than the experimental battery. The expected factor content is given, as are the number of items, the time limit, and an estimate of speededness. The word "speeded" is used to indicate that very few or no subjects finished. "Some speeding" indicates that it was estimated that twenty-five per cent to seventy-five per cent finished. Most of the tests in the experimental battery were given in two separately-timed parts in order to make possible the computation of alternate-form reliability coefficients. The coefficients corrected by the Spearman-Brown formula are given with the test descriptions. A full description of the test is not given when the name of the test or the kind of test is familiar to most readers (*2*).

For Verbal Comprehension

Verbal

(1) *Antonyms (Entrance)*. 40 items, 15 minutes. 93% finished.

(2) *Reading Comprehension (Entrance)*. 20 items, 25 minutes. 93% finished. This test was intended to measure both Verbal Comprehension and General Reasoning.

(3) *Verbal Relations (Entrance)*. Verbal-analogies items and verbal-completion items. 40 items, 30 minutes. 88% finished.

For General Reasoning

Reading Comprehension. (Described above.)

(4) *Mathematics Reasoning (Entrance)*. 20 items, 30 minutes. 61% finished. This test was intended to measure both General Reasoning and Number.

Variables 5 to 44 were scores from the experimental battery.

(5) *Figure Analogies.* Two parts, each 20 items, 10 minutes. Some speeding. $R = 0.64$.

For Space

(6) *Similar Figures.* From four squares containing black and white areas the one was to be selected which was not simply

rotated but was the mirror image of the other three. Two parts, each 40 items, three minutes. Speeded. $R = 0.89$.

(7) *Block Counting.* In pictures of large piles of similarly shaped blocks, the subjects were to count the shaded blocks that touched each numbered, unshaded block. The contacts between the blocks had to be inferred from the parts of the blocks shown. Two parts, each of five piles and 50 items, five minutes. Some speeding. $R = 0.82$.

(8) *Revised Minnesota Paper Form Board.* Series MA (Psychological Corporation). One of five patterns was to be selected which could be produced by putting together given outlined parts. 60 items, 25 minutes. Speeded.

For Speed of Judgment

(9) *Visual Preferences.* Each item consisted of a pair of somewhat similar, simple line drawings or figures. The subject checked the one he liked best. The score was the total number checked. No account was taken of the kinds of figures preferred. (For this test and the following two tests, approximately every fifth item presented a very easy choice so that subjects responding at random could be detected. No such subjects were found.) Two parts, each 50 items, two minutes. Speeded. $R = 0.88$.

(10) *Social Judgment.* Each item consisted of pairs of names of personal qualities. The subject checked the quality he would prefer to characterize someone with whom he had to associate closely. The score was the number checked. No account was taken of the kinds of qualities preferred. Two parts, each 70 items, two minutes. Speeded. $R = 0.92$.

(11) *Size Judgment.* Each item consisted of pairs of descriptions of objects. The subject checked the larger (assuming standard or average sizes). The score was the number checked. No account was taken of the correctness of the judgments. Two parts, each 50 items, two minutes. Speeded. $R = 0.92$.

For Perceptual Speed

(12) *Picture Discrimination.* Each item consisted of three simple drawings of faces, two exactly alike, and one different in some respect. The subject checked the different one in each item. Two parts, each 56 items, two minutes. Speeded. $R = 0.89$.

(13) *Cancellation.* Each part consisted of a page of typed, ran-

dom, capital letters reproduced in red. The task was to draw an X over every A, starting at the top of the page and working down. The score was the number of A's cancelled. Two parts, each two minutes. Speeded. $R = 0.96$.

(14) *Number Checking.* This test is similar to the Minnesota Clerical Aptitude Test. Two parts, each 48 items, two minutes. Speeded. $R = 0.89$.

For Number

Mathematics Reasoning (Described above.)

(15) *Checking Addition and Division.* This was an alternation of five items, each of simple addition and division of 1- and 2-digit numbers, with an answer given. The answer was to be checked as correct or wrong. Two parts, each 50 items, two minutes. Speeded. $R = 0.89$.

(16) *Subtraction and Multiplication.* This was simple subtraction and multiplication of 1- and 2-digit numbers in spiral, omnibus form with five response alternatives. Two parts, each 42 items, two minutes. Speeded. $R = 0.91$.

For Carefulness (with these variables the signs of the correlations were reversed, so that high scores would represent favorable performance)

(17) *Picture Discrimination (Wrong).* This was the wrongs score on Variable 12. $R = 0.56$.

(18) *Checking Addition and Division (Wrong).* This was the wrongs score on Variable 15. $R = 0.42$.

(19) *Subtraction and Multiplication (Wrong).* This was the wrongs score on Variable 16. $R = 0.38$.

(20) *Cancellation (Omit).* This was the number of uncancelled A's prior to the last A cancelled in Variable 13.

For Visualization

(21) *Similar Rotations.* The same solid form is shown in two different rotational positions in each item, labeled 1 and 2 respectively. A new form, labeled 3, was to be visualized as going through the same rotation in one or two planes as that represented by drawings 1 and 2. The resulting position of form 3 was to be selected from drawings A, B, C, D, or E. Two parts, each 20 items, five minutes. Some speeding: 53% finished Part 1; 41% finished Part 2. $R = 0.69$.

(22) *Folded Paper Test.* This test was similar to Thurstone's

Paper Folding Test (*2*). Two parts, each 20 items, five minutes. Some speeding. $R = 0.71$.

(23) *Mechanical Movements.* Drawings were presented of different kinds of mechanical apparatus involving gears, belts, eccentrics, etc. One or two items based on each drawing asked about the direction, speed, or nature of the movement within the apparatus pictured. Two parts, each 10 items, eight minutes. Some speeding: 25% finished Part 1; 31% finished Part 2. $R = 0.68$.

For Induction

(24) *Inducing Rules.* This test presented, in spiral omnibus form, items requiring the subject to discover a rule and then to answer the item by applying it. Included were analogies, series, and belonging items drawn from verbal, social studies, science, numerical, and graphic material. Variety was introduced to prevent the subjects from using a system for solving the problems. Two parts, each 24 items, seven minutes. $R = 0.72$.

(25) *Letter Sets.* The set of four letters not belonging with other sets of four letters was to be chosen. Two parts, each 15 items, five minutes. Considerable speeding: 7% finished Part 1; 20% finished Part 2. $R = 0.43$.

(26) *Locations.* This test was adapted from Thurstone's Marks Test. Each item presented five groups of dashes with spaces between them. An X replaced one dash in each of the first four lines, following some rule for placement. The subject was to indicate where the X should appear in line 5. Two parts, each 24 items, five minutes. Speeded. $R = 0.72$.

For Fluency of Expression

(27) *Metaphors.* This was similar to Thurstone's Similes. The subject was required to complete a simile in three different ways. The score was the number of completions regardless of their quality; no entirely senseless ones were made. Two parts, each 10 items, three minutes. $R = 0.84$.

(28) *Double Opposites.* This was similar to Thurstone's test of the same name. Each item presented a word and called for writing two words opposite to it in meaning and beginning with specified letters. The score was the total number of correct responses. Two parts, each 25 items, four minutes. $R = 0.72$.

(29) *Multiple Expressions.* Part I required the subject to write down as many ways as he could of requesting a second helping of meat. Part II required the subject to write down as many ways as he could of asking someone for his name. In each case the score was the number of expressions written down regardless of their quality; no senseless ones were made. Two parts, each four minutes. $R = 0.82$.

For Aiming or Eye-hand Coördination

(30) *Dotting.* The subject placed a dot inside as many typewriter-sized capital O's as he could. The O's were irregularly spaced along crooked lines. The score was the number of O's correctly marked. Very few were incorrectly marked. Two parts, one minute each. $R = 0.92$.

(31) *Tracing-Easy.* A line was to be drawn as rapidly as possible through lines ¼ inch apart having $\frac{3}{16}$ -inch openings. The score was the number of lines correctly passed. Very few were passed incorrectly. Two parts, one minute each. $R = 0.92$.

(32) *Tracing-Difficult.* A line was to be drawn as rapidly as possible through lines $\frac{1}{10}$ inch apart having ⅛-inch openings. The score was the number of lines correctly passed. Very few were passed incorrectly. Two parts, one minute each. $R = 0.94$.

For Motor Speed

(33) *Writing X's.* This score was the number of X's drawn on or near regularly spaced dots. Two parts, one minute each. $R = 0.92$.

(34) *Writing "lack."* This score was the number of times the word "lack" could be written (or printed) on or near regularly spaced lines. Two parts, one minute each. $R = 0.92$.

(35) *Writing Digits.* This score was the number of times the digits 1 to 9 could be written on or near given lines. Two parts, one minute each. $R = 0.91$.

For Mechanical Knowledge

(36) *Mechanical Information.* This was a multiple-choice factual information test in practical mechanics. Two parts, each 12 items, five minutes. Power conditions. $R = 0.74$.

(37) *Tool Information.* Each item presented the picture of an object used in some mechanical operation and the pictures

of three tools or other objects. The task was to select the one of the three tools which was commonly used with the first object. Two parts, each 15 items, five minutes. Power conditions. $R = 0.80$.

(38) *Mechanical Comprehension.* This was patterned after the Bennett test. Two parts, each 15 items, five minutes. Power conditions. $R = 0.64$.

Variables 39–44 were scores taken from a ninety-item multiple-choice information test designed to measure interest in six areas on the assumption that a subject who has been sufficiently interested in an area to participate in it or read much about it will have picked up factual information about it. The scores are the number of items correct among the 15 (placed in random order) representing each area. Total time, 50 minutes. Unspeeded. The areas covered were:

(39) *Interest Information (Science).* Common-sense science.

(40) *Interest Information (Social).* Games, dancing.

(41) *Interest Information (Financial).* Economics, banking, stock-market.

(42) *Interest Information (Athletic).* Baseball, sailing, bowling, etc.

(43) *Interest Information (Art and Music).* Painting, great music, technical information.

(44) *Interest Information (Mechanics).* Home shop tools, auto mechanics.

THE SUBJECTS AND THE ADMINISTRATION

The battery, five and one-half hours long including directions and administrative time, was given to about seven hundred and sixty plebes shortly before the end of their first summer at the United States Military Academy at West Point. With a ninety-minute history examination and a twenty-minute biographical inventory, it was given in two half-day sessions on successive days. In order to include some of the entrance test variables in the study, it was necessary to restrict the group to those who had been "Regular" and "Validating" candidates. This left out the "Certificate" candidates, i.e., those Congressional appointees who are certified by their schools. The number of subjects with scores on all forty-four variables was four hundred and ten.

THE CRITERION VARIABLES

Five additional variables as follows were obtained when first-year course grades became available in the summer of 1951:

(45) *Language Grades.* Choice within quotas of French, German, Portuguese, Russian, or Spanish. Fundamentals of grammar, reading, writing, and particularly oral practice.

(46) *English Grades.* Includes review of fundamentals, composition, reading of literature with class discussion and themes, and speech making.

(47) *Military Topography and Graphics Grades.* Use of instruments, engineering drawing, freehand and design sketching, map projections, descriptive geometry, and charts and graphs.

(48) *Mathematics Grades.* Includes algebra, trigonometry, solid geometry, analytic geometry, and an introduction to calculus.

(49) *Aptitude for Service Ratings.* Ratings on leadership potential. These are made by an officer advisor and by all fellow cadets in the cadet's company including those from all four classes. The officer's rating counts one-third; the cadets' ratings count two-thirds.

The product-moment intercorrelations were computed for the 44 variables on the four hundred and ten cases. For the tests which were divided into two parallel parts, the score used for the intercorrelations was the sum of the part scores.

For all correlations with the criterion variables the number of subjects was three hundred and sixty-one. This reduction was caused by normal attrition that took place during the first year at the Academy. The intercorrelations and the means and standard deviations may be obtained from ADI (*8*). Attrition at the Academy during the first year tended to raise the means and decrease the standard deviations a little, as might be expected. For tests given in two parts the alternate-form reliability coefficients corrected by the Spearman-Brown formula are given in the diagonal of the table.

THE FACTOR ANALYSIS

Factor analysis of the 49-variable intercorrelation matrix was carried out by means of Thurstone's Grouping Method (*5*). The

highest value in each column was used as an estimate of the communality.[2]

Sixteen factors were extracted before the residuals seemed to be reduced to near the chance level. The groups of variables used for the extraction of the sixteen reference factors may be obtained from ADI (*8*).

After the sixteen factors had been extracted, no pattern that seemed significant could be found among the residuals. The largest value among the residuals was 0.079. The distribution of the residuals and the unrotated factor matrix may also be obtained from ADI (*8*).

THE ROTATION

Rotation to oblique simple structure was carried out by the method described by Thurstone (*5*, p. 194), where all pairs of factors are plotted on paper and the new positions for the hyperplanes are drawn on the paper, so that the direction cosines for the new rotation may be read from the graph. Seven sets of rotations were carried out, including a total of one hundred and twenty rotated pairs. A new set of graphs was assembled after each set of rotations. The final transformation matrix may be obtained from ADI (*8*). The final oblique solution is presented in Table I. Loadings of 0.20 or higher are printed in parenthesis. Communalities derived from the unrotated matrix are given in the right-hand column. Table II shows the intercorrelations of the factors.

In the process of rotation it was borne in mind that the infinite number of rotational solutions simply represent alternate ways of explaining the intercorrelations that were found. No one solution is more mathematically correct than another. While simple structure was the general goal, this concept is so variously or vaguely defined that there is much freedom for variation within its limits. Rotation in this study was not done blind, but was carried out so that the result would be of maximum value to the broad purpose

---

[2] After the extraction of the ninth factor, the values used for the communalities were the highest values in the columns of the residual table. Up to the extraction of the ninth factor the highest values in the columns of the original table were inserted in the diagonals of the residual tables. Since no groups had overlapped up to this point, it made little difference whether the communalities of members of the groups were estimated from the original or the residual tables.

TABLE I—ROTATED FACTOR MATRIX AND COMMUNALITIES
(Decimal points omitted. Loadings of 0.20 or higher are printed in parenthesis.)

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	h^2
1 Antonyms (entrance)	(66)	−03	−04	−02	09	07	−04	−10	08	−01	−01	−02	−02	−03	04	19	79
2 Reading comprehension (entrance)	(43)	05	06	−01	−09	−07	10	06	−12	−02	−03	−01	04	08	−05	(−23)	68
3 Verbal relations (entrance)	(55)	−03	−03	04	00	01	−07	04	04	03	05	03	−02	−05	02	07	76
4 Mathematical reasoning (entrance)	04	04	−08	−03	−03	01	05	02	05	−10	09	(32)	−04	02	00	07	54
5 Figure analogies	02	04	−01	−04	09	06	−05	14	15	01	−04	18	06	12	−07	−01	52
6 Similar figures	01	−05	−01	06	−09	−03	−02	(23)	00	03	00	−06	00	(30)	04	06	45
7 Block counting	−10	−08	11	01	−01	−00	−00	(28)	−11	−07	08	03	11	(22)	−03	07	54
8 Minnesota paper form board	−10	04	18	04	−01	−03	14	(32)	00	−01	−00	−12	−03	(24)	−05	−17	57
9 Visual preferences	−12	−07	−01	−02	(59)	04	05	03	06	04	−10	−01	−01	04	01	−06	66
10 Social judgment	−01	−02	−03	02	(65)	−01	−00	−02	02	−04	−00	−07	04	02	−00	05	75
11 Size judgment	04	07	01	−01	(48)	−03	−05	−02	−05	00	18	07	−03	−05	−01	00	72
12 Picture discrimination	06	19	01	02	19	02	−05	13	05	05	−01	−07	−03	16	−01	04	65
13 Cancellation	03	(36)	02	00	−02	−02	04	−03	−18	−06	05	05	01	01	−05	05	49
14 Number checking	−03	(32)	−03	−02	03	01	−03	05	−01	04	−01	−04	−06	03	04	−07	64
15 Checking addition and division	02	−01	−01	−02	01	03	01	00	−06	01	−01	−01	−02	−01	(43)	04	85
16 Subtraction and multiplication	−00	02	−02	03	−02	−04	−02	−01	03	−01	−02	−00	02	02	(42)	−05	84
17 Picture discrimination (wrong)*	−04	02	04	−03	−05	−12	01	05	(37)	07	01	−03	−01	01	−04	(26)	41
18 Checking addition and division (wrong)*	03	05	02	−04	−01	03	03	−08	(46)	−01	−05	06	−01	−04	−04	−04	37
19 Subtraction and multiplication (wrong)*	−03	−03	−04	−03	05	07	−03	−04	(43)	−05	05	01	−08	04	03	(−22)	37
20 Cancellation (omit)*	−08	−09	01	09	01	−06	04	−01	(31)	−03	01	−00	12	−01	−01	07	23
21 Similar rotations	−06	08	02	−06	10	−01	02	−04	15	−05	05	−01	02	(31)	01	−04	40
22 Folded paper test	−04	01	07	−07	02	10	03	06	16	01	−06	11	−05	(27)	−01	−04	53
23 Mechanical movements	02	08	11	−00	03	−02	−03	−09	−01	−09	10	13	01	(28)	−03	13	60
24 Inducing rules	05	−02	−07	−03	05	03	−04	(30)	03	11	(20)	06	−03	01	06	06	24

24 Inducing rules	05	−02	−07	−03	05	03	−04	(30)	03	11	(20)	06	−03	01	−06	06	64
25 Letter sets	−03	08	06	−07	04	−02	07	(29)	−02	13	−06	08	09	−07	−03	−03	40
26 Locations	03	05	−08	−00	06	07	−03	(30)	−04	−01	−06	(21)	−06	01	−06	03	49
27 Metaphors	−11	−06	02	02	−02	00	−02	−02	−05	(56)	04	−00	01	−00	05	02	54
28 Double opposites	(25)	07	−01	−01	06	−01	−02	−09	06	(28)	01	−02	02	00	03	03	36
29 Multiple expressions	08	03	−01	−00	02	02	05	04	01	(51)	−07	03	−04	−00	−05	−05	47
30 Dotting	−06	−04	−00	09	01	(41)	02	02	03	02	03	−15	05	10	04	−01	58
31 Tracing-easy	01	07	−08	02	−02	(56)	−08	01	−04	−03	05	14	−10	−09	−06	08	70
32 Tracing-difficult	−05	−05	08	−11	−02	(60)	06	−03	02	01	−08	01	05	−02	01	−06	61
33 Writing X's	13	02	04	(47)	−04	05	−03	−03	−07	−04	−01	−01	00	−02	04	18	51
34 Writing "lack"	01	04	−02	(60)	05	01	03	−07	08	02	−01	04	−00	−01	−07	−16	68
35 Writing digits	−06	−07	−01	(68)	−01	−06	01	10	−01	03	01	−03	01	03	02	−05	71
36 Mechanical information	06	−02	(44)	−02	−01	00	−01	−09	11	−04	(29)	−02	−05	−01	−03	02	73
37 Tool information	−01	−01	(62)	01	05	−05	10	02	−01	−03	02	01	11	03	−01	01	75
38 Mechanical comprehension	06	−11	(39)	−03	−00	07	−01	06	07	08	−03	−09	02	(23)	10	−03	57
39 Interest information (science)	−04	03	02	−04	−05	02	00	−05	−01	−02	(48)	07	−04	−02	−09	01	50
40 Interest information (social)	06	−02	03	−01	01	02	−04	−07	−01	05	−04	−03	(48)	10	02	05	49
41 Interest information (financial)	08	−14	09	02	02	−07	−03	−01	−01	05	15	01	05	09	10	00	29
42 Interest information (athletic)	−06	01	−05	05	04	−03	06	09	−03	−02	04	08	(29)	−11	−07	−04	27
43 Interest information (art and music)	14	01	02	−05	−00	04	−07	04	09	−03	(25)	−18	11	03	01	−04	37
44 Interest information (mechanics)	−09	01	(43)	05	−02	−05	−02	01	02	03	(30)	05	01	−08	−09	01	63
45 Language grades	05	−04	−08	01	−02	01	(51)	04	−02	−03	−01	−03	03	−02	01	06	60
46 English grades	(28)	12	−03	−04	06	−02	(37)	02	04	06	−01	18	−04	−15	−10	02	69
47 Military topography and graphics grades	−12	−02	(24)	05	01	04	(25)	12	01	04	06	11	−05	(22)	−03	−07	74
48 Mathematics grades	−06	−01	02	01	−01	02	(45)	05	−04	−03	−03	(30)	07	04	−02	−05	71
49 Aptitude for service ratings	−06	05	(26)	−04	−00	−06	(38)	−06	−09	06	−17	04	(27)	−01	−02	−02	24

* For these variables the signs of the correlations have been reversed, so that these variables represent Carefulness.

TABLE II—INTERCORRELATIONS OF THE FACTORS

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1		−.02	.05	.01	.01	.05	.27	.35	.13	.24	.60	.46	.31	.39	.04	−.20
2	−.02		.03	.42	.38	.55	.26	.58	.16	.18	.27	.11	.33	.23	.74	.05
3	.05	.03		.05	.17	.18	−.23	.06	−.10	−.03	.24	.19	−.22	.37	−.11	.14
4	.01	.42	.05		.35	.54	.18	.13	−.07	.30	.18	.07	.11	.06	.36	−.01
5	.01	.38	.17	.35		.34	.13	.25	−.32	.49	.27	.07	.20	.08	.41	.03
6	.05	.55	.18	.54	.34		.18	.36	−.11	.19	.24	.07	.24	.27	.34	.07
7	.27	.26	−.23	.18	.13	.18		.23	.09	.15	.33	.08	−.02	.09	.34	.14
8	.35	.58	.06	.13	.25	.36	.23		.37	.16	.47	.46	.43	.48	.48	.01
9	.13	.16	−.10	−.07	−.32	−.11	.09	.37		−.28	.12	.25	.26	.15	.18	−.04
10	.24	.18	−.03	.30	.49	.19	.15	.16	−.28		.31	.13	.26	.07	.24	−.09
11	.60	.27	.24	.18	.27	.24	.33	.47	.12	.31		.51	.38	.43	.41	−.15
12	.46	.11	.19	.07	.07	.07	.08	.46	.25	.13	.51		.30	.58	.36	−.20
13	.31	.33	−.22	.11	.20	.24	−.02	.43	.26	.26	.38	.30		.08	.41	−.18
14	.39	.23	.37	.06	.08	.27	.09	.48	.15	.07	.43	.58	.08		.08	.16
15	.04	.74	−.11	.36	.41	.34	.34	.48	.18	.24	.41	.36	.41	.08		−.11
16	−.20	.05	.14	−.01	.03	.07	.14	.01	−.04	−.09	−.15	−.20	−.18	.16	−.11	

of the research, that of identifying a useful set of pure-factor tests by proving that an expected factor structure would appear. The criterion of simple structure and a positive manifold of loadings was responsible for a great majority of the decisions made in the rotation, but two examples may be mentioned to show the way in which other conditions played a part.

(1) This study does not reveal Space, Induction, and Visualization as three separate factors. Rotation could have been such that any two of these three factors would appear as primaries with the third being a combination of the other two. The final rotation was made to show the spatial tests as loading both Induction and Visualization, because this situation had the readiest psychological interpretation and, therefore, served as the best framework for the development and understanding of pure-factor tests.

(2) The second group of variables used in the extraction of factors consisted of the three "perceptual speed" tests and the two "number" tests. The first two rotations left the fifteenth reference factor as a residual, while the second was distinctly divided into two clusters, the perceptual and the number tests. The hyperplanes were then rotated so as to pass through these two clusters. Thus, Factor 2 became perceptual only, and Factor 15, which was a residual factor, became the Number factor. Although the two new factors correlate 0.74, the clear division between them provides convincing evidence that factorial invariance has occurred. These two

factors can, therefore, be relied upon to appear and may well be useful dimensions in a testing program.

INTERPRETATION OF THE FACTORS

The factors are named to indicate their identification with factors found by other investigators. Most of them are now included in the *Kit of Selected Tests for Reference Aptitude and Achievement Factors* (*4*), which was developed by committees of investigators for the purpose of defining some of the better established factors. The selection of some tests for the *Kit* was based partly on the findings of this analysis.

Following each factor name there appears the "Universal Index Number" of the factor from a numbering system for factors which was initiated by Cattell (*1*).

Because this study was undertaken to confirm the establishment of familiar factors, which have been described elsewhere (*2* and *3* and original sources referred to therein), the factor descriptions and the discussions are brief.

Factor I is *Verbal Comprehension*, Universal Index No. T13: Knowledge of the English language, including the skill necessary both to read and to write it.

Factor II is *Speed of Symbol Discrimination*, Universal Index No. T12: The ability to find a known symbol amidst a mass of distracting stimuli. Tests 12, 13, and 14 were included to define the "perceptual speed" factor. However, "perceptual speed," thought by the writer at the time of testing to be a single entity, actually consists of at least two factors, one of which is speed of symbol discrimination. Another "perceptual speed" factor calls for speed in perceiving detail. Variable 12 is a test of this other factor. Its low loading seems to confirm that fact.

Factor III is *Mechanical Knowledge*, Universal Index No. T8: A knowledge stemming from instruction or experience in the field of mechanics. This factor is unique in this battery of so-called aptitude factors, since it depends on variance in achievement. It was included in this aptitude battery because mechanical knowledge plays such an important part in our culture that it is likely to be of consequence in predicting success in various academic or vocational fields.

Factor IV is *Motor Speed*, Universal Index No. T9: Speed in

moving the fingers in complex patterns that are well fixed by habit. This may be the same as the U. S. Employment Service's Finger Dexterity factor (*2*), and seems certainly to be the same as Bechtoldt's "Motor Speed" factor (*2*), and Woodrow's "Speed" factor (*7*). It is not known to what extent this factor may be generalized beyond tasks carried out with the hand holding a pencil.

Factor V is *Speed of Judgment* (not indexed): Speed or lack of compulsiveness in making minor decisions. The large individual differences on tests of this sort are probably affected rather strongly by the situation and the test directions. However, it seems reasonable to assume that a part of this variance represents a mode of response habitual to the individual.

Factor VI is *Aiming* or eye-hand coördination, Universal Index No. T1: The three variables having high loadings on Factor VI in this study are all performed with pencil and paper. These tests occur in several studies in the literature without evidence that the factor encompasses more than the paper and pencil situation (*2*).

Factor VII is the *Grades* factor (not indexed): The variance that is common to course grades but not to any tests. This variance probably covers a wide variety of characteristics including such things as classroom attitude and parental pressure.

Factor VIII is a combination factor of *Induction* and *Space*, Universal Index Nos. T5 and T11: *Induction* involves the findings of general concepts that fit sets of data, the forming and trying out of hypotheses. *Space* is the ability to comprehend the nature of the arrangement of elements within a visual stimulus pattern primarily with reference to the location of the examinee. The zero hyperplane of this factor could have been rotated to a different position with respect to Factor XIV, but the present rotation is the one which is most meaningful. The *Space* factor appears between *Induction* and *Visualization* and lies in the plane determined by these two. This makes good psychological sense, as suggested by the term "Spatial Reasoning" so often used instead of "Space." The three variables (6, 7, and 8) put into this battery to represent the "Space" factor are shown here to be a combination of a reasoning factor and the factor called *Visualization*.

Factor IX is *Carefulness* (not indexed): The ability or disposition to proceed cautiously on speed tests. In all of these tests the nature of the task itself is easy, but the directions call for work at top speed. The table of intercorrelations indicates that there was

no consistent correlation between the carefulness scores and the speed scores derived from the same tests. Because this is true, there can be no serious distortion of the factor pattern resulting from use in the analysis of two scores derived from each of four tests.

Factor X is *Fluency of Expression* or possibly *Idea Fluency*. (The latter would be Universal Index No. T6) : The three tests were placed in the battery to define *Fluency of Expression*, the ability, once an idea is given, to express that idea rapidly in many different ways. However, the distinction between new ideas and new expressions of an idea is not clear and is not established here, because no recognized test of *Idea Fluency* was included in the battery. Such a test, for example, might call for writing down as many ideas as possible about some such topic as "a house." The score on this kind of "Idea Fluency" test would be the number of ideas written down within a given short time limit.

Factor XI appears to be *Interest in Science* (not indexed) or an information acquisition factor. Even if it is considered an interest factor, it is not known whether it, defined as it is by tests of *information* about science, is the same as that found in numerous other studies loaded by questions asking about *preference* for science. The Inducing Rules test includes questions involving a knowledge of general science, and so its loading is reasonable, but there is no explanation for the loading of Interest in Art.

Factor XII is probably *General Reasoning*, Universal Index No. T34: A factor which has been associated with arithmetic reasoning tests and a rather varied list of other reasoning measurements. Figure Analogies was included in the battery to help define this factor. Its loading of 0.18 is lower than expected, but it is the highest loading for Figure Analogies on any of the factors.

Factor XIII is an interest or personality factor. It could be called *Interest in People* or it may be similar to the *Sociability* factor found in personality inventory analyses. Factors having both of these interpretations have been found in many studies (*3*).

Factor XIV is a combination of *Visualization* and *Space*, Universal Index Nos. T14 and T11: As noted for Factor VIII, the inference here is that the tests used for the "Space" factor actually require the abilities represented by both *Visualization* and *Induction. Visualization* may be defined as the ability to comprehend objects in space so as to manipulate them in the imagination.

Factor XV is the *Number* factor, Universal Index No. T10: Speed in carrying out simple numerical computations involving the four arithmetical processes. Mathematics reasoning tests ordinarily have a loading on this factor as well as on *General Reasoning*. In this analysis the mathematics variables do not load on *Number*, because *Number* and *General Reasoning* were rotated to a correlation of 0.36 with each other.

Factor XVI has no apparent psychological meaning. It can best be considered a residual factor.

CONCLUSION

Tests of each of thirteen well-established factors and some tests of general information were administered to an entering class at West Point. Together with scores from the entrance examinations and first-year grades, these were factor-analyzed to confirm the identification of useful pure-factor tests by finding factorial invariance when such tests are factored.

All thirteen of the sets of tests used in the battery for the designation of reference factors did in fact turn out to fall into clusters within the factor space. Twelve out of thirteen of the clusters defined factors as expected. Those three variables intended to represent *Space* did not load a separate factor for themselves but did form a thirteenth cluster located in the plane defined by *Induction* and *Visualization*. The *General Reasoning* factor is included among the twelve that were separated successfully. However, the Mathematics Test (No. 4) and Figure Analogies (No. 5) do not define the factor well by themselves. The factor does become adequately clear, however, because of the presence of Mathematics Grades (No. 48).

The other three of the fifteen non-residual factors make good psychological sense: (1) The *Grades* factor has often appeared in studies where several course grades and tests were included. (2) The *Interest in Science* factor is one that has been found several times in the literature (*3*). Leadership ratings and amount of information on social activities combine to define a factor. The implication here is that the leadership ratings are heavily weighted with sociability.

From the standpoint of the study's purpose, it may be concluded that factorial invariance held up, and, therefore, the selection of tests to represent certain factors was successful. The clusters are

sufficiently tight to make it reasonable to represent factors in a battery by one test from each set. Thus, the study served as a pre-test yielding factorial, reliability, and level-of-difficulty data upon which to base the selection of tests.

REFERENCES

1) R. B. Cattell, A universal index for psychological factors. Advance Publication No. 3, Laboratory of Personality Assessments and Group Behavior, Univ. of Illinois, Urbana, Ill., 1953.

2) J. W. French, "The Description of Aptitude and Achievement Tests in Terms of Rotated Factors," *Psychol. Monogr.*, No. 5, 1951.

3) J. W. French, *The Description of Personality Measurements in Terms of Rotated Factors.* Princeton, Educational Testing Service, 1953.

4) J. W. French, (Ed.), *Kit of Selected Tests for Reference Aptitude and Achievement Factors.* Princeton, Educational Testing Service, 1954.

5) L. L. Thurstone, *Multiple-Factor Analysis.* Chicago, Univ. of Chicago Press, 1947.

6) Ledyard R. Tucker, "A Semi-Analytical Method of Factorial Rotation to Simple Structure," *Psychometrika,* 9: 43–68, 1944.

7) H. Woodrow, "The Relation Between Abilities and Improvement with Practice," *J. educ. Psychol.,* 29: 215–230, 1938.

8) American Documentation Institute. A more detailed form of this paper (or extended version, or material supplementary to this article) has been deposited as Document number 5163 with the ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington 25, D. C. A copy may be secured by citing the Document number and by remitting $1.25 for photoprints, or $1.25 for 35 mm. microfilm. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress.

RETENTION OF GENERAL PSYCHOLOGY

W. J. McKEACHIE and DANIEL SOLOMON

University of Michigan

Whenever a group of teachers gather to commiserate about the futility of trying to educate the average college student, someone is sure to say, "Within a week after the final examination the students will have forgotten three-quarters of what they've learned anyway!"

Such statements, based on Ebbinghaus's classic studies of memory have seldom been tested in terms of the amount of gain in knowledge during a course retained over a period of time following the course.[1] Ebbinghaus, himself, discovered that meaningfulness made for improved retention. We might, then, expect retention to vary somewhat from course to course as well as individual to individual.

If an instructor is teaching an advanced course, how much should he expect students to remember from previous courses? This study had the practical objective of determining how much of students' learning in a course in general psychology was retained over a summer or more.

Procedure

At the University of Michigan a course in general psychology is a prerequisite for the course in educational psychology. In the fall of 1954 most of the students in the educational psychology course had taken the general psychology course either a year earlier or the preceding spring. Accordingly, a forty-item multiple-choice test was constructed from items on the general psychology final examinations for those two semesters. This test was given to students in the educational psychology course during the first

[1] Most studies of retention of learning in school have simply compared scores on an examination at the end of a course with scores on the same test at some later date. Such a procedure ignores the possibilities that students might score as well on many or all of the test items at the beginning of the course as they do at the end of the course. The model which seems appropriate is illustrated by the study of Frutchey (*1*) who found in a study of high school chemistry that general principles are retained better than specific terminology.

TABLE I.—PERCENTAGE OF CORRECT ANSWERS TO EXAMINATION ITEMS

	Pre-test	End of General Psychology Course	Beginning of Educational Psychology Course Sept., 1954	Per cent Retention
19 items common to June, 1954, and Educational Psychology tests	54.03	73.5	70.7	85.6
18 items common to Feb., 1954, and Educational Psychology tests	41.12	78.0	71.1	81.3

week of the fall semester.[2] Thirty-five of the students entering the educational psychology course had also taken the final the previous June. There were nineteen items common to the two tests. Eighty-one students had taken the final the previous February. There were eighteen questions common to these two tests.

In June the mean score on the nineteen common items was 14.0. In September the same students averaged 13.4. While test papers were not available for the February examination, item analysis data gathered for all students in the course indicated that the mean on the eighteen common items had been 14.0, while the mean score of eighty-one of these students in educational psychology eight months later was 12.8. Since students entering education are on the average no better than the average student in our course, this too indicates a high degree of retention.

This was a much higher degree of retention than we had anticipated. We began to wonder whether or not our results could be explained by the hypothesis that our test scores were not really measuring course achievement. Perhaps the students would have done as well on these items even without taking the course. To check this we gave the same test to a group of students at the beginning of the course. The results showed their scores to be significantly lower than those of the educational psychology students.

As Table I indicates, if we take the gain in test score from beginning to the end of the course as measure of learning, retention over a four to 8-month period is more than eighty per

[2] The authors appreciate the help of Dr. Robert Dixon in arranging for administration of this test. Charles Smith and Diedrich Snoek assisted in test scoring and analysis.

cent. While the possibility remains that the test measures some general gain in sophistication independent of the course, it seems fairly safe to conclude that material learned in the first course in psychology is retained for an appreciable period after the final examination.

REFERENCE

1) F. P. Frutchey, "Retention in High-School Chemistry," *Journal of Higher Education,* 8: 217–218, 1937.

ASSESSMENT OF INTELLIGENCE LEVEL BY THE RORSCHACH TEST AS COMPARED WITH OBJECTIVE TESTS

BOLESLAW A. WYSOCKI

Marquette University[1]

Rorschach wrote that his test " . . . presents an intelligence test almost completely independent of previous knowledge, memory, practice and degree of education" (*5*, p. 183). This high estimate of blots as to assessing intellectual abilities is shared, almost unopposed, by most of the Rorschach experts. They contend that intelligence is one of the most important factors that can be measured with a great accuracy by the Rorschach test. Thus the purpose of this paper will be to find out to what extent the group method of administration proves or disproves Rorschach's well-known contention concerning the assessment of intelligence level by his ink blots.

TESTS

The following tests were used in the present study: (1) Group Rorschach, scored after Klopfer (*2*, *3*, *4*); (2) Verbal Intelligence Test, "S.P. Test 15," used by the British Army and consisting of one hundred equally weighted items; (3) Nonverbal Intelligence test, Raven's Matrices, consisting of sixty equally weighted items.

SUBJECTS

In order to make comparisons with Rorschach categories, the verbal and nonverbal intelligence tests were given to one hundred and thirty-two men and eighty-five women. The ages ranged from eighteen to forty-six with an average of 27.9 years for men, and from sixteen to forty-three with an average of 21.5 years for women. The average educational background was grade 10.3 and 9.6 for men and women respectively.

[1] Now at the Alliance College, Cambridge Springs, Pa.

TABLE I.—CORRELATION COEFFICIENTS BETWEEN RORSCHACH CATEGORIES AND INTELLIGENCE TESTS

Rorschach Categories	Verbal Intelligence Test		Nonverbal Intelligence Test	
	Men	Women	Men	Women
No. of R	0.440	0.389	0.389	0.450
W%	0.239	0.331	0.329	0.112
M%	0.272	0.368	0.266	0.389
F + %	0.356	0.111	0.129	0.111
F − %	−0.115	−0.174	−0.083	−0.053
A%	−0.291	−0.205	−0.124	−0.184

RESULTS

In order to establish validity, if any, of the group Rorschach method in assessing intelligence level, the several Rorschach categories were correlated with the two intelligence tests.[2] In all cases coefficients were calculated by the full product-moment formula (Pearson) and are shown in Table I.

Correlations between the several Rorschach categories and the intelligence tests of 0.228 and 0.283 are significant at 0.01 level for men and women respectively; whereas correlations of 0.174 and 0.217 are significant at 0.05 level for men and women respectively.

As Table I shows the highest positive correlation with intelligence tests, significant at 0.01 level (both the verbal and nonverbal) is N of R, that is, total number of responses for men and women. The second category showing positive correlation at 0.01 level of significance with both tests is W, a complete whole response. Next to the W is the M category (human movement), showing positive correlations at 0.01 level of significance with both tests for men and women. Good form F + %, gives only one positive coefficient significant at 0.01 level with the verbal intelligence test for the male group. The poor form, F − %, gives negative correlations not significant at 0.01 or 0.05 levels. Animal responses, A, did correlate significantly with the intelligence tests at 0.01 level for men and 0.05 level for women.

Next, the possibility of combining the scores in the various Rorschach categories was taken into consideration to see if a

[2] Coefficient of correlation between the verbal and nonverbal intelligence tests was 0.707.

TABLE II.—INTERCORRELATIONS OF INTELLIGENCE TESTS AND RORSCHACH CATEGORIES

	Verbal Int. Test	No. of R	W%	M%	F+%	F−%	A%
N. verbal Int. Test	0.707	0.424	0.273	0.203	0.193	−0.052	−0.191
Verbal Int. Test	—	0.409	0.290	0.179	0.209	−0.146	−0.229
No. of R		—	0.384	−0.083	−0.054	−0.067	−0.222
W%			—	.092	0.064	0.093	−0.196
M%				—	0.334	−0.140	−0.209
F + %					—	−0.503	0.087
F − %						—	0.044
A%							—

combination would be better than the best single score. For this purpose the Wherry-Doolittle test selection method was adopted (*1*, p. 435) to satisfy need. Thus, the matrix of intercorrelations between the two intelligence tests and six Rorschach possible indicators of intelligence was computed. The correlations are simple product-moment coefficients for the whole group of subjects, that is, for one hundred and thirty-two men and eighty-five women tested in the study. The correlation matrix is given in Table II. Correlations of 0.181 are significant at 0.01 level and correlations of 0.138 are significant at 0.05 level.

Next, corrected multiple correlations were calculated as follows:

Rv.1	=0.409	Rn.1	=0.424
			(R of 0.181 is significant at 0.01 level)
Rv.14	=0.464	Rn.13	=0.492
			(R of 0.212 is significant at 0.01 level)
Rv.146	=0.487	Rn.134	=0.499
			(R of 0.234 is significant at 0.01 level)
Rv.1463	=0.487	Rn.1345	=0.503
			(R of 0.253 is significant at 0.01 level)
Rv.14632	=0.490	Rn.13456	=0.505
			(R of 0.269 is significant at 0.01 level)
Rv.146325	=0.487	Rn.134562	=0.505
			(R of 0.283 is significant at 0.01 level)

In the above-listed multiple correlations 1 stands for No. of R; 2 for W%; 3 for M%; 4 for F+%; 5 for F− %; and 6 for A%. The symbol v represents the verbal intelligence test, and n the nonverbal intelligence test.

From the review of these multiple correlations one may observe

that by adding several Rorschach categories to the first one selected (with the highest correlations with intelligence tests) substantial gains are obtained increasing multiple R from 0.424 to 0.505 and from 0.409 to 0.490 for intelligence tests, both nonverbal and verbal, respectively. On the basis of the above multiple correlations it may be noted that, when taken together, four Rorschach categories, namely number of responses, human movement, good form and animal responses, constitute the best combination of intelligence predictors, as measured by intelligence tests.

SUMMARY AND CONCLUSIONS

An attempt was made to determine the validity of the group Rorschach in assessing the level of intelligence. In the first stage of this investigation, the group Rorschach results were compared with two intelligence tests, verbal and nonverbal, for men and women separately. Significant correlations, ranging from 0.272 to 0.450, were found for the following Rorschach categories: total number of responses, the percentage of whole responses, the percentage of human movement responses, good form per cent, and the percentage of animal responses. The correlations between the poor form level and intelligence tests were found not significant when calculated for men and women separately.

Further, a consideration was given to the possibility of exploring combinations of several scores instead of single ones. Thus, after computing the correlation matrix for men and women together, several corrected multiple correlations were calculated. It appeared that by successive additions of single Rorschach scores the multiple correlation was significantly increased up to 0.081 R points. It also turned out that the following Rorschach categories constituted the best combination of scores for predicting level of intelligence: number of responses, human movement, good form, and animal responses.

These results confirm those of Wishner (*6*) with regard to the individual Rorschach.

REFERENCES

1) H. E. Garrett, *Statistics in Psychology and Education*. New York: Longmans, Green and Co., 1947.

2) M. R. Harrower-Erickson and M. E. Steiner, *Large Scale Rorschach Techniques*. Springfield, Illinois: Charles C. Thomas, 1945.

3) B. Klopfer, et al. *Developments in the Rorschach Technique*. New York, Yonkers-on-Hudson, 1954.

4) B. Klopfer, *The Rorschach Technique*. New York, World Book Co., 1942.

5) H. Rorschach, *Psychodiagnostics* (translated from German by P. Lemkau and B. Kronenberg). Bern, H. Huber, 1942.

6) J. Wishner, "Rorschach Intellectual Indicators in Neurotics," *American Journal of Orthopsychiatry*, 18: 265–279, 1948.

BOOK REVIEWS

SAMUEL J. BECK. *The Six Schizophrenias.* New York: American Ortho-Psychiatric Association, 1954, pp. 237, plus viii.

An important contribution to Rorschach researches is Dr. Beck's volume on the six schizophrenias. The study purports to be, and is, a coördinated inquiry into the problem of schizophrenia as carried on by the three scientific disciplines—psychology, social work and psychiatry. This is the orthopsychiatric pattern for the multi-discipline approach in the study of human behavior. The researches, one with children and one with adults, were conducted in 1947–52 in the Michael Reese Hospital in Chicago. The aim of the research group has been to describe the schizophrenic person by operational trait data which are clinically and psychodynamically meaningful and to inspect these descriptions by sound method.

Six patterns of dynamic organization of the traits emerged and the volume is chiefly concerned with a description of details of the study and the results found.

Dr. Ray R. Grinker contributes the initial chapter called "Clinical Introduction" and in it he describes the step-by-step build up of the research and the clinical significance of it for research and therapy. Considered in terms of methods and approaches, the research includes: (1) The analysis of a sizeable series of normal adults and children—sixty children, fifty adults; (2) an analysis of the clinical histories and the Rorschach tests of non-schizophrenics; (3) an analysis of clinical history and study of clinical formulations with Rorschach correlates.

The Stephenson Q technique used for sorting of clinical data and the corresponding Rorschach data on each case is considered in a special chapter included concerning Q methodology by Prof. Stephenson.

It is Dr. Grinker's opinion that the researchers give us not only a better idea of dynamic process of schizophrenia, a better way of understanding the process of disintegration and attempts at reintegration, a better predictive of diagnostic view of schizophrenia, but also a much more satisfactory way of viewing the personality. Dr. Beck perceives the researches as having four implications: (1) Describes the schizophrenia as certain over-all dynamic ad-

justment patterns; (2) emphasizes latency in schizophrenia that is not necessarily psychosis; (3) bimodality; (4) postulate of least resistance. And he is happy about the fact that the test as a tool has withstood the reactions of the tough minded group of psychiatrists that helped in his research.

There is no questioning the value of these researches for stimulating further research, as well as adding to our knowledge concerning the Rorschach as a tool in the studying of personality. What can be questioned is the practicality of the approach in terms of making the most useful contribution towards developing insights for understanding and treatment of schizophrenia.

However, the frame of reference for the researches is understandable enough considering Beck's persistent interest in the Rorschach test.

H. MELTZER
Human Relations Research Foundation, St. Louis, Missouri

DELOS D. WICKENS and DONALD R. MEYER. *Psychology.* New York: Dryden Press, 1955, pp. 541 plus vi. $5.25.

IRWIN A. BERG. *Student's Workbook.* New York: Dryden Press, 1955, pp. 181 plus v.

HARRY P. BABRICK. *Teacher's Manual.* New York: Dryden Press, 1955, pp. 70.

Another text in Psychology is this book by Wickens and Meyer. In it the authors endeavor to present the essential facts about behavior as they are now known and also a systematic method of explaining which they consider the most fruitful, rigorous and promising. As far as intellectual orientation is concerned, the chief emphasis in this book is on the importance of learning and the use of learning concepts in the analysis of complex forms of behavior. The early chapters of the book deal with learning concepts and are captioned "Learning Fixation" and "Utilization of Learning." These are followed with chapters dealing with basic concepts in areas of motivation and perception, social behavior and personality development, dealt with in the light of the basic learning concepts as presented in the earlier chapters. The last few chapters deal with physiological foundations of behavior and a very brief last chapter represents an attempt to sum up. Each chapter is followed by a list of references. The book contains an

index. The publishers have done an excellent job in printing and preparing a book of attractive appearance.

This new psychology text includes the usual chapters, though somewhat rearranged. Diagrams and pictures indicate a fair amount of resourcefulness used in putting the volume together to make ideas considered clear. References are relatively of a more recent origin. Behavior is described by the authors as "responses made to stimuli acting on the individual at the particular moment. Because of the individual's biological structure and the nature of his past experience, these stimuli will evoke certain psychological mechanisms—such as habits and motives, to mention only two—and certain kinds of responses will follow." This approach, the authors say, does not deny conscious experience, it simply states it will not use conscious experience as the main source of explanation. In short, what they are saying is that they are dominantly behavioristic.

Academic newcomers in the field are likely to be less critical of the general content and emphasis. People who have grown with the field are likely to miss certain experiments, illustrations and conclusions that they are used to seeing and thinking of as being fundamental in presentation of the subject of psychology. People who are psychoanalytically oriented will not find the content or emphasis in this book satisfying. Their treatment of "learning to learn," therapists and clinicians will find lacking in an emphasis on need for unlearning. In the authors' consideration of forgetting, they cover qualitative experiments in forgetting. Qualitative in the sense of learning and forgetting more meaningful life experiences is not mentioned by them. Homeostasis is used in a completely uncritical manner. Persons who think that it is well for students to know something about historical foundations of the field, including some of the founders, will find the book disappointing in this respect. Thorndike's name, for example, is not mentioned and some of the other leaders in the field are barely mentioned. Though the writing is relatively well done, the emphasis on experimental data gives certain chapters popular interest, a feeling of being dried up.

What the authors have tried to do in this book, to a fair extent they have achieved in doing. A fair amount of material usually included in a first course is treated here compactly and in a relatively simple manner. Chapters on physiological foundations at the

end of the book are relatively well done, but the kind of material students expect to find in psychology books is treated in less humanized fashion than needs be. All in all they have done well what they consciously strived to do. And omissions and shifts of interest obviously are intentional. It is a safe book to use for a psychology text rather than the most advisable one.

Student's Workbook for use with Psychology by Wickens and Meyer is prepared by Irwin A. Berg. It includes a listing of concepts, terms and objective test exercises of each chapter. It also contains some demonstrations, exercises and formal statements for demonstrations and projects.

Teacher's Manual to accompany the Wickens and Meyer book is prepared by Harry P. Babrick of Ohio Wesleyan University. It includes a list of text questions and discussion questions for use by teachers. It also contains an annotated bibliography of recent literature relevant to the teaching of general psychology.

H. MELTZER
Human Relations Research Foundation, St. Louis, Missouri

MILTON GREENBLATT, RICHARD H. YORK, and ESTHER LUCILE BROWN. *From Custodial to Therapeutic Patient Care in Mental Hospitals.* New York: Russell Sage Foundation, 1955, $5.00, pp. 497.

"From Custodial to Therapeutic Patient Care in Mental Hospitals" is a publication of the Russell Sage Foundation, which was established in 1907 "for the improvement of social and living conditions in the United States of America." The psychiatrist on the team, Dr. Milton Greenblatt, writes the chapters on "Towards Therapeutic Community" and with the social worker, E. Brown, writes the conclusions. The psychologist, Richard York, writes Part II, which is concerned with experiments in ward patient care in Veterans Administration Hospitals, and Part III where the problem of motivating patients on chronic wards at Metropolitan State Hospital is considered. The book also contains an excellent bibliography on social treatment compiled by Dr. Frederick L. Wells. It also contains an index.

That mental hospitals are behind times is no news for people in the field and for an increasing number of other people. The hospitals whose function it is to administer to the sick minds,

rather than to the sick bodies, of nearly three quarters of a million patients, continue to struggle with buildings of which some are fire traps and almost all are overcrowded. They struggle also with an inadequate staff and inadequate facilities of all kinds. Care is still largely in the hands of attendants of whom there are supposedly from 80,000 to 90,000 and many of these positions are held by men and women who have neither the needed aptitude nor experience. The purpose of this volume is not to tell a weeping story of neglect for mental hospitals, but to present a description of work done and work that can be done to reduce the present confusion and discouragement. It is basically an exploratory review in social treatment. The attitude of the authors is eclectic in nature. As they see it, the function of a mental hospital is to utilize every form of treatment available for restoring patients to health. And at worst, to aid them to live as nearly normal lives as possible within an institutional setting. The goal of the authors is to use both the personnel and physical resources of the hospital to aid in the development of competence in the patients. The description of ward and ward treatment included will be of interest to all people working with institutional cases or in institutions. It will be of particular interest and use to administrators of mental hospitals. In their assessments of hospitals, collaboration with community sources and resources, the authors show a nice balance between the sense of limitations and the continued planning for improvement. In their emphasis on the need for improvements of conditions in chronic wards, the authors are calling attention to a phase that at the present time still is perhaps too often neglected. The authors are perfectly willing to consider that another kind of social environment completely outside the hospital could be devised which is better suited to the needs of some chronic patients. Their hope is that from studies (of the kind reported in this volume) of work done on the wards might evolve decisions concerning greater differentiation of care and the opportunity, as least for the mental hospital to become a hospital in fact, as well as in name.

One thing is certain—in all the mental hospitals which have been modernized, particularly by the addition of a better trained staff, an encouraging number of patients with acute conditions have been discharged. And thanks to help from the Federal Government, as well as increased awareness on the part of many

state governments, sights are being raised and at least the beginnings of promise are in view. The question facing us now, as described by the authors, is "whether hospitals will strive to maximize the use of such physical and social resources as exist, including resources in patients, to achieve this end; and whether the American public will become a more active partner in this venture by greatly increasing contributions and sustaining interest and service."

The description of the material here, as well as the attitude expressed should be of help to all people connected with or working with mental patients.

H. MELTZER
Human Relations Research Foundation, St. Louis, Missouri

C. WINFIELD SCOTT and CLYDE M. HILL. *Public Education Under Criticism.* New York: Prentice-Hall, 1954, $4.75, pp. 414 plus xiv.

HENRY EHLERS. *Crucial Issues in Education.* New York: Henry Holt & Co., 1955, $2.25, pp. 277.

To the wave of criticism of our public schools and education in general, there is only one answer—more and better education. And that includes education of educators and interested laymen to an awareness of the meaning and significance of the issues. By so doing, the editors of the two anthologies on education hope to contribute toward the rescue of America from the retreat from learning advocated by critics. The two volumes being reviewed are both concerned with the problem of how best to handle the criticism of education and how to develop realistic attitudes for action to improve the educational scene.

"Public Education Under Criticism" represents an attempt on the part of professional educators to consider the needs and the best ways of handling the criticisms of teachers, school administrators and teaching techniques. It makes an honest effort to let opponents and proponents present their particular attitudes for the purpose of appraising the credentials of our educational philosophy as well as practices. Selections made are presented under five sections, considered as areas of criticism and ways of handling them. The major criticisms include progressive education, fundamentals, religion, social studies, teacher education and teachers

and general defenses. The section following is concerned with analysis of these criticisms as well as the critics, and this section, in turn, is followed by one on the handling of criticisms. A relatively smaller last section is concerned with concluding statements which list generalizations and suggestions for constructive action. Professional educators, as well as enlightened lay people, will find much of interest and value in this anthology.

The general conclusion is that most of the criticisms are made by honest, perhaps well-intentioned, even though often misinformed or uninformed people. Also that the most grossly exaggerated criticisms come from prestiged, though uninformed, people, people without experience with schools. And that the group of opponents of public education most to be feared are the vicious people who have baneful reasons for opposing public education. Their opposition to schools, the authors view as a part of a subversive movement to destroy all democratic institutions. The articles are presented in a double column arrangement. The print is small, though reading fairly easy. The authors are many and most of the articles relatively short.

The volume called "Crucial Issues in Education" is more concerned with a level of consideration for the intelligent layman, rather than the educator. In this volume are considered the pros and cons of educational issues mostly from the writings of the past ten years. The editorial viewpoint is expressed in the introduction, which reads, "True education should open many channels of inquiry; it should consider many types of belief. Lest education be reduced to propaganda, teachers must not only be informed about the many partisan groups and interests around them, but must also be courteous and considerate toward various conflicting viewpoints."

In this anthology, the authors make an attempt courageously to face the issues and rather than running away from issues that are controversial in nature, they aim to meet up with them first, then consider them and make decisions about them. The materials have been selected and the authors chosen with this in mind.

The material selected is presented under such captions as Freedom for Teachers, Freedom for Learners, Religion and Public Education, Racial Segregation and Education, Classroom Method and Materials, and a concluding chapter on Democracy and Education. Illustrative of contributors included are Judge William O.

Douglas, Learned Hand, Howard Mumford Jones, Justice Black, Henry Steele Commager. Each chapter is followed by questions and readings for study and discussion.

In both of these volumes there is included much material of an intelligible nature that can serve people well who want to know about the waves of criticism about education today, what the viewpoints are and the likely background for these viewpoints and what the implications are for education in democracy.

H. MELTZER
Human Relations Research Foundation, St. Louis, Missouri

PUBLICATIONS RECEIVED

GEORGIA SACHS ADAMS and THEODORE L. TORGERSON. *Measurement and Evaluation for the Secondary-School Teacher.* New York: The Dryden Press, 1956, pp. 658. $5.75.

AKE BJERSTEDT. Interpretations of Sociometric Choice Status. Lund, Sweden: C. W. K. Gleerup, 1956, pp. 408 (paper). 40 sw. cr.

EDGAR F. BORGATTA and HENRY J. MEYER, Editors. *Sociological Theory: Present-Day Sociology from the Past.* New York: Alfred A. Knopf, Inc., 1956, pp. 546. $6.75.

DANIEL BROWER and LAWRENCE E. ABT, Editors. *Progress in Clinical Psychology. Volume II.* New York: Grune & Stratton, 1956, pp. 364. $7.75.

HARRISON BULLOCK. *Helping the Non-Reading Pupil in the Secondary School.* New York: Bureau of Publications, Teachers College, Columbia University, 1956, pp. 180. $3.75

D. DEE CASTETTER, LLOYD S. STANDLEE and NICHOLAS A. FATTU. *Teacher Effectiveness: An Annotated Bibliography.* Bloomington, Indiana: Institute of Educational Research, School of Education, Indiana University, 1954, pp. 105 (paper). $1.00.

The College Entrance Examination Board. Fifty-fourth Report of the Director. Princeton, New Jersey: College Entrance Board, 1955, pp. 107 (paper). $0.50.

Contributi del Laboratorio di Psicologia. Milano, Italia: Universita Cattolica del Sacro Cuore, Piazza S. Ambrogio, 9, 1955, pp. 544 (paper).

DAVID G. DANSKIN. *Workbook in Educational Psychology.* New York: Henry Holt and Co., 1956, pp. 249. $2.75.

RAY DOTTERER. *Postulates and Implications.* New York: Philosophical Library, Inc., 1955, pp. 509. $7.50.

ALLEN L. EDWARDS. *Statistical Methods for the Behavioral Science.* New York: Rinehart & Company, Inc., 1954, pp. 542. $6.50.

VERGILIUS FERM, Editor. *Encyclopedia of Morals.* New York: Philosophical Library, Inc., 1956, pp. 682. $10.00.

FRIEDA FROMM-REICHMANN and J. L. MORENO. *Progress in Psychotherapy.* New York: Grune & Stratton, Inc., 1956, pp. 352. $8.50.

J. RAYMOND GERBERICH. *Specimen Objective Test Items: A Guide to Achievement Test Construction.* New York: Longmans, Green and Company, Inc., 1956, pp. 436. $4.75.

LEONARD V. GORDON. *Gordon Personal Profile and Gordon Personal Inventory.* Yonkers-on Hudson: World Book Company, 1956 (paper). $0.50.

ANDREW W. HALPIN. *The Leadership Behavior of School Superintendents.* Columbus, Ohio: The Ohio State University Press, 1956, pp. 109 (paper). $2.00.

WILLIAM E. HENRY. *The Analysis of Fantasy: The Thematic Apperception Technique in the Study of Personality.* New York: John Wiley & Sons, Inc., 1956, pp. 305. $6.00.

WILLLIAM A. HUNT. *The Clinical Psychologist.* Springfield, Illinois: Charles C. Thomas, 1956, pp. 207. $5.50.

A. M. JORDAN. *Educational Psychology. Growth and Learning.* Fourth Edition. New York: Henry Holt and Co., 1956, pp. 600. $5.50.

MAURICE H. KROUT, Editor. *Psychology, Psychiatry, and the Public Interest.* Minneapolis: The University of Minnesota Press, 1956, pp. 217. $4.00.

HENRY H. LINN, Editor. *School Business Administration.* New York: The Ronald Press Company, 1956, pp. 574. $7.50.

RICHARD MADDEN and PHILIP PEAK. *Madden-Peak Arithmetic Computation Test.* Yonkers-on-Hudson, New York: World Book Company, 1956, pp. 18 (paper).

STANLEY S. MARZOLF. *Psychological Diagnosis and Counseling in the Schools.* New York: Henry Holt and Co., 1956, pp. 401. $4.00.

A. DA SILVA MELLO. *Man: His Life, His Education, His Happiness.* New York: Philosophical Library, 1956, pp. 729. $6.00.

HENRY ESTILL MOORE. *Nine Help Themselves.* Austin, Texas: The University of Texas, 1955, pp. 286. (paper).

E. A. PEEL. *The Psychological Basis of Education.* New York: Philosophical Library, Inc., 1956, pp. 303. $6.00.

SANDOR RADO and GEORGE E. DANIELS, Editors. *Changing Concepts of Psychoanalytic Medicine.* New York: Grune & Stratton, Inc., 1956, pp. 248. $6.75.

ANNE ROE. *The Psychology of Occupations.* New York: John Wiley & Sons, Inc., 1956, pp. 340. $6.75.

DAGOBERT D. RUNES. *On the Nature of Man; An Essay in Primitive Philosophy.* New York: Philosophical Library, Inc., 1956, pp. 105. $3.00.

MORRIS S. SCHWARTZ and EMMY LANNING SHOCKLEY. *A Study in Interpersonal Relations.* New York: Russell Sage Foundation, 1956, pp. 289. $3.50.

F. ELEANOR SCHONELL. *Educating Spastic Children.* New York: Philosophical Library, Inc., 1956, pp. 242. $6.00.

GEORGE D. SPACHE and PAUL C. BERG. *The Art of Efficient Reading.* New York: The Macmillan Company, 1955, pp. 273 (paper). $3.00.

CATHERINE E. STELTZ. *As You See It: Twenty-four Pictures to Stimulate Group Discussion of Social and Moral Problems, with Suggestions for Their Use.* New York: Bureau of Publications, Teachers College, Columbia University, 1956, pp. 26 (paper). $2.95.

DAVID A. STEWART. *Preface to Empathy.* New York: Philosophical Library, Inc., 1956, pp. 157. $3.75.

PERCIVAL M. SYMONDS. *Dynamics of Psychotherapy: The Psychology of Personality Change. Volume I. Principles.* New York: Grune & Stratton, Inc., 1956, pp. 211. $5.50.

W. D. WALL. *Education and Mental Health: A Report Based upon the Work of a European Conference Called by UNESCO at the Musee Pedagogique in Paris.* New York: Columbia University Press, 1955, pp. 347. $3.00.

ROBERT W. WHITE. *The Abnormal Personality.* Second Edition. New York: The Ronald Press Company, 1956, pp. 644. $6.50.

J. WAYNE WRIGHTSTONE, JOSEPH JUSTMAN, MORRIS PINCUS and RUTH H. LOWE. *New York Test of Arithmetical Meanings. Level One and Two.* Yonkers-on-Hudson, New York: World Book Company, 1956, 48 pp. (paper).

(Concluded from inside front cover)

these charges, as well as information regarding the services of the American Documentation Institute in making available tables, etc., that for economic reasons may be omitted from a printed article, by writing to the Editors or the Publishers.

Double-spacing.—Manuscripts should be typed, written on one side of the paper only, and double-spaced throughout including quotations, footnotes, and bibliographical references.

Footnotes.—Footnotes are to be numbered consecutively beginning with '1', and should be on a separate sheet at the end of manuscript. (Footnotes to tables carry the *, †, and ‡.)

Titles.—Titles of articles should be brief, preferably three to eight words, with an extreme maximum of twelve words.

Type style.—Manuscripts are not to be marked for type style—this is done in the editorial office.

Books and other materials for review, and business communications should be addressed to THE JOURNAL OF EDUCATIONAL PSYCHOLOGY, Warwick & York, Publishers, 10 E. Centre St., Baltimore 2, Md.

Subscribers should notify the Publishers of change of address at least four weeks in advance of publication of the issue with which change is to take effect; both the old and new address should be given.

The Publisher desires every subscriber to get all issues to which he is entitled. Each journal is securely enclosed in a sturdy wrapper on which the subscriber's name and address have been printed, and is delivered directly to the Post Office, postage prepaid. Second-class matter is handled less by postal employees than other mail; moreover, if the Post Office is unable to make delivery, a notice to this effect is sent the Publisher and the magazine returned. Consequently, it is doubtful if one journal in many thousands is actually lost in transit.

But after an issue has been delivered to the proper address many things may happen to it—it may be diverted, or misplaced, or borrowed and not returned. For this neither Post Office nor Publisher is responsible. However, a subscriber who does not find a given issue in its assigned place may innocently make a claim of non-receipt. No claim for non-receipt of an issue can be honored unless made within four weeks after arrival of the next succeeding number. In order that a claim may arrive within the time limit it should be addressed to the Publisher—not to an agency.

WARWICK AND YORK *Publishers* BALTIMORE 2, MD.